Self-Regulation and Reinterpretation in the Nietzschean Self

Syea, Emma

Awarding institution:
King's College London

Self-Regulation and Reinterpretation in the Nietzschean Self

Emma Syea

A Thesis submitted for the degree of

Doctor of Philosophy

February 2018

Abstract of the Thesis

This thesis examines Nietzsche's model of subjectivity, with a particular focus on the process of self-regulation. Self-regulation is to be understood as the maintenance of the psychological stability or 'well-being' of an individual, especially in the face of adverse circumstances. I suggest that self-regulation provides us with a more intelligible means of exploring the philosophical psychology Nietzsche provides than self-formation. The reason for this is twofold: (i) self-regulation is a phenomenon exhibited by all individuals, not just elite, 'higher' types, who are singled out by Nietzsche, as being capable of the task of self-formation, and (ii) we can identify distinct psychological mechanisms involved in this process of self-regulation, in contrast to the somewhat obscure notion of 'becoming what you are'. The thesis explores key psychological mechanisms involved in regulating a self, namely forgetting, autobiographical memory, and self-deception. The capacities implicated in these regulatory processes can be seen to possess a different functional profile to that of a drive or an affect. The thesis thus supplements current drive-based accounts which have so far dominated discussions of Nietzsche's model of the self, by further exploring different structures within the self, and their operations. Outlining the interactions between these regulatory capacities, and the drives and affects, enables us to have a richer understanding of how a self may develop itself through interpretation, incorporation, and extirpation. Finally, the thesis offers a taxonomy of the characteristics of self-regulation by triangulating this notion with Nietzsche's concept of health. This will show how self-regulation can be linked with marks of Nietzschean health such as incorporation, providing a necessary condition for self-formation in strong, artistic, 'higher' types. However, the thesis will also show that self-regulation can come apart from health, concluding that self-regulation emerges as a necessary but not sufficient condition for the higher ideal of health.

Acknowledgements

To paraphrase a friend's wise advice: 'no PhD thesis was ever written by just sitting alone at a desk, scribbling down some notes. Unless you're Kant.' I can confirm that, from my experience at least, this certainly rings true.

Thanks first and foremost to my supervisor Sacha Golob. I could not have asked for a better supervisor and mentor. He has supervised me through an undergraduate degree (dealing with hapless essays written on Kantian metaphysics!), a masters degree, and now the PhD, going above and beyond what is required from a supervisor at each stage. Sacha has taught me so much more than how to grapple with Nietzsche's thought: I have learnt what it means to be a philosopher and a teacher. I cannot thank you enough.

I am very grateful to my second supervisor Michael Silk for countless inspiring conversations about Nietzsche, and for advice and encouragement throughout the project.

Thank you to Eliot Michaelson who worked on the project with me for a year. Eliot helped me to find the puzzles in Nietzsche, and made trying to solve them less daunting.

I have benefited enormously from conversations and written feedback from the following people: Manuel Dries, Fabrice Teroni, Tom Stern, Andrew Huddleston and Ken Gemes. Thank you all for providing interesting and invaluable perspectives on my work.

I have been fortunate enough to be a member of a welcoming and collaborate department, something which has made all the difference during this process. Thanks to Clare Moriarty, Paul Doody, Matt Bodle, John Wright, and Mike Coxhead in particular.

Thank you to Rod and Sonja Nash who have not stopped encouraging me, and showing an interest in my academic career since my very first German lesson at school, aged twelve. Introducing me to the German language opened the door to a wealth of literature and philosophy, including, of course, Nietzsche.

I could not have done the PhD without good friends around me – you have encouraged, supported, and generally put up with me through the ups and downs!

Thanks to: The good doctor Charlie Bell, my Nietzscheans Matthew Dennis, Sander Werkhoven, and Lorenzo Serini, and my Kantian Paola Romero, my departmental partner in crime Vanessa Brassey, and my Epistemologist extraordinaire Giada Fratantonio, and to all the usual suspects: Nick Rutter, Elizabeth Gourd, Emma Obertelli, Anvita Sharma, Alison McGarvey, Christine Good, Sam Scarpa, Alex Clarke, Laura Silva, Ashleigh Ellis, Sam Shetabi, and Adam Dickson. You have all provided much needed moments of sanity and laughter. I am lucky to have you all.

And thank you to Ian Jakeway for the adventures.

Finally, thank you to my parents for always supporting me and never doubting me.

Table of Contents

List of Abbreviations

A	*The Anti-Christ*
BAW	*Frühe Schriften*
BGE	*Beyond Good and Evil*
BT	*The Birth of Tragedy*
D	*Daybreak*
EH	*Ecce Homo*
GM	*On the Genealogy of Morality*
GS	*The Gay Science*
HAH	*Human, all too Human*
KSA	*Kritische Studienausgabe*
KGW	*Werke Kritische Gesamtausgabe*
TI	*Twilight of the Idols*
TSZ	*Thus Spoke Zarathustra*
UM	*Untimely Meditations*
WC	*The Wagner Case*
WP	*The Will to Power*

Where reference to the *Kritische Studienausgabe* is made, the volume number is used, followed by the note number, e.g. KSA 9, 11 [141].

Where works are composed of individual essays, following the abbreviation is the essay number (Roman numeral) and the section number (Arabic numeral), e.g. GM II:10 refers to *On the Genealogy of Morality,* the second essay, section ten.

In works that include titled main sections, the section title is used, followed by subsection numbers (where applicable), e.g. EH 'Why I am so Wise' 2 refers to *Ecce Homo,* the main section 'Why I am so Wise', subsection two.

Introduction

i. The Project

When examining Nietzsche's understanding of subjectivity, commentators have tended to view Nietzsche's stance on the self as going hand in hand with his notion of self-formation or 'becoming what you are'. This thesis will instead examine Nietzsche's model of subjectivity via an alternative process in the self, namely self-regulation, and suggest that the activity of self-regulation provides us with a more intelligible means of exploring Nietzsche's philosophical psychology than self-formation.

The reason for shifting the focus from self-formation to self-regulation is twofold. Firstly, the exact nature of self-formation is fraught. There is no clear consensus on what this process consists in and the accounts can vary widely; May (2009) appeals to drive psychology, and takes self-formation to be a unification of the multiplicity of drives by a master drive within the self, whilst Nehamas (1985) understands the process in highly aestheticized terms, where the individual 'blends' his actions into a perfectly coherent whole via the medium of literature. Often the notion of self-formation is tied up with ideas of freedom and agency, (notions which are themselves somewhat contested). Some commentators such as Leiter (2001), doubt the very possibility of self-formation, claiming that it is not consistent with Nietzsche's characterisation of drives as fixed. Others, such as Katsafanas (2012), have tried to allay this concern by drawing upon the idea that consciousness, as distinct from the drives, and causally efficacious, can open up the possibility of self-formation. Self-formation is also often aligned with ideas of self-knowledge or self-understanding, the status of which appears ambiguous in Nietzsche's writings. The very notion of self-formation thus appears deeply mysterious, and it is not clear how deeply this process is meant to operate, i.e. whether self-formation amounts to self-styling in the sense of taking up second-order attitudes against one's character, or whether more concrete changes take place at the level of the drives.

Secondly, self-formation is reserved for the few. Often characterised as an achievement or an ideal, Nietzsche envisages this phenomenon as open only to the 'higher' men: artistic types who have a specific psychological constitution, e.g. Goethe is one of the few singled out as having achieved this ideal (TI 'Skirmishes of an Untimely Man' 49). It is not an option for everyone. Trying to understand the self via a notion of such exclusivity may not teach us very much about the range of operations available in the philosophical psychology Nietzsche outlines. Focusing on self-formation may be restrictive in this sense.

Self-regulation, by contrast, is something which every individual exhibits (albeit in varying degrees of efficiency). How should we understand self-regulation? This notion can have various connotations, e.g. self-control, self-managing, self-monitoring, and it can be applied to spheres as diverse as behaviour, impulses, emotions, thoughts, and health. Within the context of Nietzsche's thought, self-regulation is often aligned with the idea of self-control relating to one's drives, and the Sovereign Individual is sometimes seen as a model of this kind of self-regulation through the notion of promise-keeping, responsibility, and commitment. However, commentators such as Acampora (2006) raise concerns about attributing too much weight to the Sovereign Individual, suggesting that this notion is too idealistic and Kantian in flavour, with its associated concepts of autonomy and responsibility.

To avoid these concerns, I instead take self-regulation to be broadly construed as the maintenance of the individual's psychological stability or 'well-being', especially in the face of challenging circumstances. This process thus relates specifically to the individual's mental economy. Self-regulation, understood this way, inevitably brings up the notion of health, a concept which preoccupied Nietzsche throughout his writings. After outlining various means of self-regulation available in the Nietzschean self, the thesis concludes by arguing that Nietzsche's concept of health emerges as a higher ideal, and that these two notions can come apart. Self-regulation can be seen as a necessary but not sufficient condition for health.

I will suggest that self-regulation can be achieved in a number of ways, and that it can be found in both, by Nietzsche's lights, healthy (e.g. the nobles in *On the Genealogy of Morality,* and Nietzsche himself in *Ecce Homo*) and unhealthy individuals (e.g. the slaves

in the *Genealogy*). We will see how self-regulation can be linked with other marks of Nietzschean health such as incorporation and reinterpretation, and that in such cases where self-regulation occurs (i.e. in strong, healthy, artistic types), it provides a necessary condition for self-formation (where this notion is understood as involving interpretation). In this way, before even addressing the notion of self-formation, we need to first investigate the prior and fundamental process of self-regulation.

Examining the process of self-regulation will shed light on different structures within a Nietzschean self. With regards to the content of a Nietzschean self, the drives and affects have tended to be the main focus of commentators, e.g. Leiter 2002, 2007, Gemes 2009, Risse 2005, who characterise these as the basic explanatory items in Nietzsche's philosophical psychology. Though some commentators allow for other cognitive capacities (e.g. the intellect or consciousness, which may interact with the drives and affects), there has not been a sustained discussion of the capacities explored in this thesis, namely, forgetting, autobiographical memory, and imagination. I suggest that these capacities possess a different functional profile to that of a drive or an affect, and tap into the different kinds of attitudes a self can exhibit, which include, as Janaway (2009:52) suggests: 'not only the standard beliefs and desires of current-day moral psychology, but also "wills", feelings, sensations, moods, imaginings, memories, valuations, convictions, and more'. Without further exploration of these capacities, and how they interact with the drives and affects, we will not have a full picture of a Nietzschean self, and accounts may appear too thin in relation to Nietzsche's practical philosophy. We will see how these methods of self-regulation have important implications for other Nietzschean notions such as Amor Fati, the Eternal Return, and Nietzsche's call for reinterpretation. In this way, self-regulation used as a prism to explore Nietzsche's views on the self, can help bridge the gap (cf. Gardner 2009) between his theoretical and practical philosophy.

ii. Methodological Preliminaries

With regards to methodology, Nietzsche's works have been consulted in the original, alongside established translations. When attempting to reconstruct arguments from the texts, Nietzsche's manipulation of the German language should not be overlooked: more

often than not, bringing out different nuances of a word will affect the interpretation of an entire passage. Where appropriate, I have amended the translation for key terms.

The canonical status of the *Nachlass* is obviously fraught. In light of this, I have, for the most part, stayed within the published works. I believe that the most forceful and considered arguments or remarks made by Nietzsche on the topic of the self are to be found within the published works. Moreover, there are more than enough substantive passages to examine and interpret within these works alone. When constructing my own arguments and interpretations, I have been keen to avoid attributing any undue explanatory or argumentative weight to fragments from the *Nachlass*, and I have sought to engage with commentators who also place the bulk of their discussion upon the published works.

Nietzsche is a notoriously unsystematic thinker, something which has tended to split commentators regarding how best to approach his writings. Preserving historical accuracy and contextualisation on the one hand, with extracting philosophical insights and conceptual coherence on the other, can be a tricky balance to maintain.

I have endeavoured to combine close exegetical reading, with occasional use of contemporary analytic philosophy as a tool for illuminating aspects of Nietzsche's philosophy. I have found that the latter can provide a useful prism through which to contest, develop, and reconstruct elements of Nietzsche's thought. Given the focus of this thesis, not to make use of relevant material from contemporary debates in philosophy of mind, philosophy of emotion, and epistemology to enrich and develop my interpretations, seemed amiss. In doing so, however, I have been careful to avoid either attributing anachronistic readings to Nietzsche, or attempts to 'shoe-horn' him into contemporary debates. Rather, I hope to have offered an exploratory project into Nietzsche's insights regarding subjectivity and self-regulation, with the aid of some analytic conceptual tools.

iii. Thesis Outline

The structure of the thesis is as follows. I motivate the project by examining how commentators' treatment of the Nietzschean self has tended to revolve around the issue of self-formation, and the shortcomings of this approach, before proposing that self-

regulation may provide a better framework for examining Nietzsche's remarks on the self. I then examine three methods of self-regulation within the self, namely forgetting, autobiographical memory, and self-deception, the capacities involved, and the effects these mechanisms have within the subject's mental economy. In this way, I explore components of Nietzsche's philosophical psychology other than the drives and the affects, and examine how these capacities interact with the drives and affects. The final chapter brings together the three previous chapters in offering a characterisation of self-regulation in general, exploring how this differs from Nietzsche's concept of health, and how it provides the necessary psychological conditions for health.

Chapter One offers a review of contemporary debates in analytic philosophy regarding the status of the self in Nietzsche. I discuss the recent trend of interpreting Nietzsche as a naturalistic thinker (Leiter 2002, 2007, 2012), and the advantages and disadvantages of such an approach. I then examine the transcendental approach (Gardner 2009a), cognitivist accounts (Katsafanas 2005, 2012, 2013, and Janaway 2009, 2012), and accounts which focus on the idea of a self as an achievement (Gemes 2009a and Anderson 2012). I evaluate the extent to which these accounts provide a comprehensive analysis of Nietzsche's views on the self, and examine what motivates the debate for each of these commentators, be it normativity, agency, intentionality, or consciousness.

I suggest that the debate can appear unclear, and often conflates conceptions of the self. I argue that this is largely due to: (i) commentators working with differing interpretations of key notions like free will, or consciousness, and (ii) the fact that each model assumes a different understanding of self-formation.

I suggest that focusing on self-formation is a somewhat problematic way of approaching the debate, since it is not clear what Nietzsche had in mind with this notion, and it is a process which is limited to a few individuals. As such, it may not provide us with an exhaustive analysis of Nietzsche's views on the self. These concerns motivate the need for a new approach: I suggest that the concept of self-regulation may be fruitful here, and that this calls for an examination of structures within the philosophical psychology provided by Nietzsche other than the drives and affects, such as memory, forgetting, and imagination.

Chapter Two examines in detail Nietzsche's account of active forgetting, and argues that it is a key regulatory procedure within the self, conducive to Nietzsche's understanding of health. Nietzsche presents what seems like an idiosyncratic understanding of forgetting. Contrary to the notion of forgetting as a passive, degenerative process, associated with decline and incapacity, i.e. poor memory performance, Nietzsche offers us a far more nuanced, and more importantly positive, understanding of forgetting. Drawing on recent studies in cognitive science, in particular the theory of 'motivated forgetting' (Anderson 2014), I suggest that Nietzsche's understanding of forgetting is best understood as an operative subconscious capacity (akin to, for example, a language faculty, i.e. functioning but beneath our conscious control). Prima facie the idea of active forgetting appears paradoxical, but when it is understood as a capacity, we can see that Nietzsche provides us with a philosophical psychology which does not in fact meet this paradox.

I then explore the structure and operations of this capacity, and how it has a different functional profile to that of a drive. I argue that forgetting does not fit the criteria of a drive outlined by commentators such as Anderson (2012), Katsafanas (2005, 2012, 2016) and Janaway (2009). I suggest that this forgetting capacity is responsive to, and recruited by the drives in order to reach their ends, forming subordinate relational structures. I then examine the most important role forgetting has within the self as a regulatory 'deep-clean' procedure, and why this process is required. I also discuss how the Nietzschean account of forgetting importantly differs from the Freudian model. Forgetting frees up space in the mental economy via alterations of the affective valence of memories. Such space is vital for the re-interpretative process and allows new values to displace old ones. I also explore the importance of this revised model of Nietzschean forgetting to the understanding of several other Nietzschean themes such as the Eternal Recurrence and Amor Fati, both key components of his practical philosophy.

Chapter Three explores the role of autobiographical memory in the self. Autobiographical memory can be understood as a type of memory which relates to an individual's personal past, and which necessarily carries a first-person perspective. Within the context of a Nietzschean self, autobiographical memory can be seen as a plastic and creative capacity that is responsive to, and works alongside drives. Using *Ecce Homo*

as a case-study, I propose that autobiographical memory emerges as an important factor in constituting a narrative self for Nietzsche, and goes some way to allowing us to locate the sense of first-person subjectivity required for such a self. Although *Ecce Homo* has often been viewed as a highly aestheticized instance of self-formation, I will use it solely as an example of autobiographical memory as it functions broadly.

I argue that the use of autobiographical memory, and the ensuing narrative in *Ecce Homo* allows for the following: 1) an awareness of a self persisting throughout time, 2) a degree of self-knowledge regarding past perspectives, and 3) a sense of first-person perspective which is tied to affectivity.

In terms of self-regulation, a narrative self allows us to make sense of the question 'who am I?', a question which preoccupies Nietzsche in *Ecce Homo.* The idea of a narrative self provides us with a degree of coherence or unity, a means of evaluating ourselves, and (in healthy individuals) a means of incorporating experiences in such a way that, through the interactions between autobiographical memory and the drives, our self-esteem is safeguarded, as a celebratory, triumphant narrative emerges.

In this way, *Ecce Homo* emerges as a good example of how the self can recuperate and regulate itself through this use of autobiographical memory

Chapter Four then turns to an example of 'unhealthy' self-regulation found in the *Genealogy*, namely the self-deception of the slaves. I begin by offering a taxonomy of the self-deceptive beliefs present in the *Genealogy,* drawing a distinction between the slaves' value-facing beliefs ('I believe that humility is intrinsically good'), and their person-facing beliefs ('I am humble'), and exploring how these beliefs can be maintained.

I suggest that the greatest potential threat to the slaves' self-deception is their affective experiences, namely their emotions such as vengefulness, hatred, anger (and the particular intensity/frequency/duration of these emotions), which run the risk of undermining their self-deceptive person-facing beliefs. In this way, it is possible that the subject's phenomenology alone may call their self-deceptive beliefs into question, as it would be seen to directly undermine their beliefs about their moral character. This kind of tension would need to be alleviated in some way in order to allow the self-deception to be preserved.

I argue that the conflict between the slaves' self-deceptive person-facing beliefs, and their high-intensity affective experiences is not accommodated for by deflationary accounts of self-deception e.g. Mele 2000. I suggest that Nietzsche gestures towards one possible means of resolving this tension between affective states and person-facing beliefs, namely via the mechanism of pretense.

I outline what constitutes a pretense-state in the *Genealogy*, and how they are to be differentiated from belief-states. Pretense here is broadly understood as an imagined-state, which plays an intermediary role between the subject's affective states and their beliefs. I suggest that these pretense-states modify the subject's affective states, e.g. vengefulness, such that they are brought into line with the subject's self-deceptive beliefs. Modification of these affective states, however, should not be understood as an actual transformation in kind of the emotion, but rather as a second-order reinterpretation of it.

The pretense-states in the *Genealogy* take the form of rich, vivid narratives (cf. GM I:15). I argue that because of their narrative richness and pretense's ability to generate affective responses (cf. Moran 1994, Gendler 2007), these pretense-states are more effective than, for example, additional belief-states in easing possible cognitive dissonance. The imaginary simulation of, e.g. revenge present in these narratives provides the slaves with something close to fictional emotions. This enables the slaves, in an acceptable context, to actively engage with, rather than repress or remove emotions of hatred, anger, vengefulness. The narratives allow the slaves to view their emotions as justified, i.e. these are the emotions experienced by 'the virtuous' and 'the blessed' who have been wronged.

Self-deception (understood as involving pretense-states) emerges as a crucial means of self-regulation for the slaves in the face of adverse circumstances, allowing them to reinterpret their phenomenology.

Chapter Five aims to identify the key features of the self-regulation process in a Nietzschean self, and argues that self-regulation emerges as a necessary but not sufficient condition for Nietzsche's distinctive understanding of health. I argue against defining self-regulation as strength or flourishing, and suggest that whilst it initially appears to be more easily aligned with health, these two notions in fact, also come apart. By tri-

angulating self-regulation with health, I argue that we can come to understand self-regulation as an independent notion, whilst also providing a more fine-grained understanding of Nietzschean health.

I argue against defining health by reducing it to other Nietzschean concepts like overcoming (Letteri 1990), or interpretation (Neuhouser forthcoming). This approach is too broad, and it is often unclear what work the concept 'health' is serving in these accounts. This approach also admits cases as 'healthy' which Nietzsche himself emphasizes are sick. I suggest that this is a result of focusing on passages where Nietzsche discusses health in abstract terms, and that we ought to shift our focus onto the distinct case-studies of health and sickness which Nietzsche provides, for example the nobles and slaves in the *Genealogy.* I then explore the consequences of an alternative approach to defining health, namely considering the normative dimension to formal and dynamic drive accounts of health (Huddleston 2017). I suggest that, whilst this approach is generally helpful in explaining why individuals are classified as sick or healthy in Nietzschean terms, there are certain cases in which it appears that the health of an individual is somewhat fixed, and that the option of amelioration or recovery from sickness is ruled out.

Using case-studies from the *Genealogy*, I then demonstrate how self-regulation can come apart from health. I discuss: (i) the possibility of self-regulation in sickness, using the example of the slaves, and (ii) an instance of fragile self-regulation in health as seen in the nobles, identifying the key mechanisms involved in both cases. I then offer a taxonomy of the self-regulation process, outlining general methods, namely interpretation, incorporation, and extirpation, and the key characteristics which emerge, namely stability/endurance/resilience, productivity and functional efficiency, adaptivity, and physiological/psychological integrity. The fact that efficient self-regulation can occur independently of health goes some way to explain why Nietzsche presents the sick as the 'greatest danger to man' (GM III:14). Finally, I suggest that the idea that self-regulation is a necessary but not sufficient condition for health is of some assistance in illuminating Nietzsche's notion of life.

1. Nietzsche and The Self

Introduction

The nature of the self is of central importance in Nietzsche's philosophy: it is a theme to which he returns throughout his writings, and it serves as the bedrock for aspects of his thought such as his critique of morality. As such, Nietzsche's views on the self and subjectivity have been the focus of much debate in recent scholarship. This chapter offers a review of how analytic commentators have approached this debate. Much of the discussion has revolved around whether, and how, a self is able to engage in self-formation. At times, this question is addressed directly, at others it is implicit in the treatment of Nietzsche's remarks. It is not clear however, that there is any general consensus regarding what self-formation in fact amounts to, and this is reflected in the variety of ways in which Nietzsche's remarks have been interpreted. As a result, the debate can appear unclear, and often conflates different conceptions of selfhood.

This chapter will suggest that these confusions in current scholarship arise partly due to a tendency to frame the investigation with this question of self-formation. This is problematic for two reasons: a) it is not clear exactly what Nietzsche had in mind with the notion of self-formation, and b) self-formation is not an option for all individuals: Nietzsche only singles out 'higher' men as capable of such an activity, cf. Goethe in TI 'Skirmishes' 49. Due to the exclusivity of self-formation, viewing the debate through this particular lens may not yield an exhaustive account of Nietzsche's model of a self. In addition, the majority of the accounts examined in this chapter restrict their discussion to the drives and affects (though some allow for other cognitive capacities). As such, we may not be able to fully account for processes such as reinterpretation, which Nietzsche takes to be part of what it means to be a self. It therefore seems that we need to open up the discussion to involve other components of the philosophical psychology Nietzsche provides, and that we need to find a different way of approaching this investigation, namely via the notion of self-regulation.

Preliminaries

Before assessing the debate, it is worth considering whether this enterprise of modelling a Nietzsche self is feasible, and if so, the kind of considerations which should be kept in mind.

Various models of the self have been put forward, which purport to be representative of Nietzsche's views on the self across his entire oeuvre. It is not obvious, however, that a single, unified view on this subject is to be found in his works. The opinions Nietzsche espouses in his earlier works such as *Daybreak,* and later works such as *Beyond Good and Evil,* are not always consistent. His emerging views on the self depend greatly upon his stance regarding, for example, free will, causal determinism, compatibilism, incompatibilism, many of which appear to be modified in later works (discussed by Katsafanas: 2012). It is necessary to take these modifications into account when attempting to construct a model of the self. We must therefore ask whether it is even possible to extract a coherent, consistent model of the self from Nietzsche's works. As Williams (1994:66) notes, Nietzsche's writings are often 'booby-trapped not only against recovering theory from it, but, in many cases, against any systematic exegesis that assimilates it to theory.' Stern (2015) even argues that Nietzsche is far more concerned with attacking prevailing models of the self than constructing his own positive model.

Alongside changing opinions over time, we also find contradictions regarding topics such as free will (sometimes even within a single text), but these can often be resolved or explained by identifying a qualified interpretation of these topics, i.e. Nietzsche is targeting a specific conception of free will over other conceptions. It will therefore be important to identify which specific conceptions of the following Nietzsche targets: 'I', the 'soul', 'free will', 'agency', 'responsibility', 'consciousness', 'autonomy', 'unity'.

The models discussed in this chapter appear to be motivated by very different concerns. There is little consensus on what it is for a self to be a self *per se*, let alone what it is to be a Nietzschean self. Gardner (2010), for example, takes the debate to be about normativity; for Katsafanas (2012) it is to do with conscious willing; whilst for Anderson (2012), the issue is the 'I'. For Gemes (2009a), and Janaway (2012), both of whom suggest that Nietzsche does not intend to seriously engage in metaphysical speculation, the emphasis

is on a self qua task or achievement. Due to the fact that these interpretations have differing concerns, at times the debate can appear unclear, and conflates conceptions of the self and/or self-formation.

The models of the self discussed here are all placed against the background of the naturalism debate in Nietzsche. Leiter's influential *Nietzsche on Morality* (2002) seeks to show Nietzsche's unwavering commitments to naturalism, and regards any departures from the naturalistic interpretation as trivial. In this way, Leiter takes very seriously Nietzsche's claim in BGE 230 to 'translate man back into nature'.

Leiter is left with an unreservedly fatalistic reading of Nietzsche with little room for notions like creativity, agency, autonomy, or responsibility. Should we wish to integrate a Nietzschean self into a normative framework, it seems that the naturalistic reading will be somewhat unattractive. Many of the alternative models here are keen to avoid this interpretation, and attempt to show that a strict naturalistic reading of Nietzsche, aside from being unsatisfying regarding practical and axiological contexts, does not cohere with many of the passages in Nietzsche's works where he criticises the 'clumsy naturalists' (BGE 12), and is unable to allow for any notion of self-formation.

In examining the various models, I will discuss the following: a) what exactly each interpretation takes the debate to be about, b) the extent to which they offer an exhaustive analysis of the philosophical psychology Nietzsche provides, and c) what they have to say about the idea of self-formation.

The first section will examine two opposing accounts: the naturalistic account espoused by Leiter (2002), and the transcendentalist account put forward by Gardner (2009).

1.1 The Naturalistic Account (Leiter 2002)

At one end of the spectrum, Leiter's naturalistic reading of Nietzsche offers a view of the self where self-formation is effectively ruled out. Leiter (2002:8) attributes to Nietzsche a 'Doctrine of Types', whereby 'each person has a fixed psycho-physical constitution which defines him as a particular type of person'. On this reading, the self is viewed

merely as a bundle of drives and affects: 'soul as social structure of the drives and affects' (BGE 12). There is nothing above and beyond these components. The individual's conscious life is reduced to what Leiter calls 'type-facts', i.e. facts about the physiology and unconscious psychology of human beings. These type facts are, according to Leiter's interpretation (2001:287), 'causally primary in fixing the trajectory of that person's life.'

This has several notable consequences. Firstly, the causal primacy of the type-facts ensures that free will does not exist for us. We are fully determined beings at the mercy of our type-facts which are fixed from birth. Here, Leiter takes 'free will' to be free will *causa sui*, unconstrained by causal factors – an incompatibilist view. Qua conscious self or 'agent', the individual does not play an active role.[1]

Instead the individual is 'an arena in which the struggle of drives (type-facts) is played out; how they play out determines what he believes, what he values, what he becomes' (2002:100). Leiter refers to a passage from *Daybreak* to reinforce this interpretation:

> [*T*]*hat* one *wants* to combat the vehemence of a drive at all, however, does not stand within our own power; nor does the choice of any particular method; nor does the success or failure of this method. What is clearly the case is that in this entire procedure our intellect is only the blind instrument of *another drive*, which is a *rival* of the drive whose vehemence is tormenting us. ... While "we" believe we are complaining about the vehemence of a drive, at bottom it is one drive *which is complaining about the other*; that is to say: for us to become aware that we are suffering from the *vehemence* of a drive presupposes the existence of another equally vehement or even more vehement drive, and that a *struggle* is in prospect in which our intellect is going to have to take sides.
>
> (D 109)

Secondly, consciousness is not causally effective and is reduced to mere epiphenomenon - a 'phantom'.

[1] Leiter does not seem to consider other conceptions of free will such as compatibilist views or the idea that we may possess free will in a limited sense in so far as our lives have various possible trajectories but we have some autonomy regarding which trajectory we take.

> The 'inner world' is full of illusions and phantasms: will is one of them. The will does not do anything anymore, and so it does not explain anything anymore either - it just accompanies processes, but it can be absent as well. The so-called 'motive': another error. Just a surface phenomenon of consciousness, an 'after-the-fact' that hides the *antecedentia* of an act more than it reveals them... What follows from this? There are no mental causes whatsoever.
>
> (TI 'The Four Great Errors' 3)

Leiter draws our attention to Nietzsche's claim that 'a thought comes when 'it' wishes, and not when 'I' wish' (BGE 17), to further vindicate his argument for the epiphenomenality of consciousness. Here, it is made clear that consciousness has no causal autonomy: 'while a person's conscious states may be part of the causal chain leading up to action, they play that role only in virtue of type-facts about the person' (2002:92). We may believe an action to have been caused by some conscious state or other, but in reality the action has been caused by some non-conscious state, be it a drive or a physiological state. Leiter therefore holds that it is a mistake to 'conceive of ourselves as exercising our will' (2007:7). In this way, something like valuing cannot be classed as an instance of cognising: there is no rational connection with cognitive activity, instead all we experience is an expression of a drive or physiological fact. Values then are explained psychologically or physiologically via antecedent, non-normative, efficient causal conditions.

Thirdly, the self is not sufficiently transparent for us to make any claims regarding the motives of the agent's actions. Actions are unknowable since 'nothing…can be more incomplete than [one's] image of the totality of *drives* which constitute [a man's] being' (D 119). We simply do not have epistemic access to what the causally effective motives really are. As a result of this epistemic opacity, we cannot assess actions in terms of their motives.

It is worth at this point examining exactly how Leiter conceives of consciousness. We may think of consciousness in the following ways: 1) awakeness/awareness, 2) reflection, 3) self-consciousness, or 4) conscious willing. Leiter, in his account, does not discuss these various nuances. Instead, he assumes that consciousness amounts to conscious willing, glossing the 'inner world' in terms of the will. It is *this* sense of consciousness which Leiter

argues is epiphenomenal. He views the content of experience solely due to the function of the drives, as opposed to conscious willing: in this sense, our actions are not products of consciousness. However, it is not obvious that should we define consciousness in a different way, e.g. awareness, that Leiter would be able to claim that it is epiphenomenal. Consciousness qua awareness or awakeness appears as a necessary condition for experiencing the world, and one which is not necessarily epiphenomenal. On this definition, consciousness is a passive feature of the self: it plays an information role as opposed to a directional role, since it does not feature as a motivational factor in agency.

As discussed above, Leiter draws on the following quotation from Nietzsche to substantiate his argument against conscious willing: 'a thought comes when 'it' wishes, not when 'I' wish'. However, as Katsafanas (2005) notes, this quotation is taken out of context:

> A thought comes when 'it' wishes, not when 'I' wish, so that it is a falsification of the facts of the case to say that the subject 'I' is the condition of the predicate 'think'. It thinks; but that this 'it' is precisely the famous old 'Ego' is, to put it mildly, only a supposition, an assertion, and assuredly not an 'immediate certainty'. After all, one has even gone too far with this 'it thinks'—even the 'it' contains an interpretation of the process and does not belong to the process itself.
>
> (BGE 17)

If we examine the entire passage it appears that:

> Nietzsche is not concerned, here, with the question of whether conscious states are causally efficacious in their own right. Rather, he is attacking the idea that there is a substantive faculty, an Ego, standing behind conscious thoughts and generating them *ex nihilo.*
>
> (2005:11)

Katsafanas goes on to argue that Leiter equivocates between: 1) attacks on consciousness as a faculty, and 2) attacks on consciousness as a property of mental states. We should not, however, assume that the denial of consciousness qua faculty, i.e. qua Ego which produces conscious thoughts *ex nihilo*, is equivalent to the claim that conscious mental

states are epiphenomenal. This is further supported by the fact that Nietzsche seems to, as Katsafanas suggests, offer an account of conscious states which does not depend upon a commitment to such a faculty.

Leiter's interpretation leaves us with a self that is static, reductive, fatalistic, and passive: a picture which seems at odds with the idea of 'becoming what you are' and self-formation. One feature of this account which particularly undermines the possibility of self-formation, is the status of conscious thought. Leiter's characterisation of conscious thought seems to leave no room for any meaningful engagement with normativity, something which might be required for the idea of self-formation.

Gardner (2010), crystallises this concern by discussing the relation between conscious thought, judgement, and legislation. He argues that the naturalistic interpretation encounters difficulties regarding the self-understanding of the Nietzschean subject: what exactly does such a subject understand itself to be doing when it legislates? Gardner views legislation as the curation of values, whereby the individual can change their configuration of values in response to rational consideration.

On Leiter's picture, however, legislation occurs outside of any framework of rational norms. Any valuation is regarded as the expression of the causal effects of psychological or physiological facts about the individual. The subject is therefore expected to recognise that there is no rational warrant for taking valuation to be anything more than an expression of these pre-normative natural forces, and that the subject qua conscious agent plays no role in legislation.

Given that Nietzsche repeatedly encourages the creation, and innovation of values in his writings, the naturalistic model seems deeply unsatisfying. Revaluating and creating value is a fundamentally discriminatory activity, in that it is the postulating of one thing as better than another. This requires judgement which in turn requires some kind of rational reflective activity. Such an activity would surely incur disagreement as to which values are better than others. In response to this, Leiter claims that, for Nietzsche, the world is 'value-less', and that the above problem will only arise if we think there are objective values. In this way, Leiter advances an anti-realist interpretation of Nietzsche

on value, i.e. nothing has value in itself. On this reading, it would be impossible to sustain any disagreement as regards which values are better than others.

Leiter draws our attention to a passage in *The Wagner Case*, where Nietzsche describes 'noble' and Christian morality as 'opposing forms in the optics of value [*Werthe*]', and asserts that, as opposites they 'are…immune to reasons and refutations. One cannot refute Christianity; one cannot refute a disease of the eye…. The concepts 'true' and 'untrue' have, as it seems to me, no meaning in optics' (WC Epilogue). Here, Leiter (2013), claims that Nietzsche is suggesting:

> there may be rational grounds for thinking one view better than another, perhaps for thinking one true and the other false, but since reasoning has so little impact in this context, it is "meaningless" (in the sense of pointless) to raise issues of truth and falsity.

In other words, Leiter takes Nietzsche to be claiming that evaluative judgements are immune to the effects of reasoning, and that instead judgements about values are just expressions of drives. 'Higher types' or 'free spirits' will have no issue with understanding that 'this is good' has no ground beyond their 'deeming it to be good' (Leiter: 2010).

However, Leiter's response is somewhat unsatisfactory. To say that for Nietzsche the world is 'value-less' seems too crude a claim. We may concede that Nietzsche argues against values conceived of as eternal, transcendental and objective qua Platonist/Christian conceptions of values, but throughout his works he is careful to elevate *Tugenden*– virtues.[2] *Tugend* for Nietzsche is closely related to the Greek αρετή - *arete* (excellence or virtue), and is presented as something which is achievable by certain types. In a notebook of 1888 Nietzsche outlines a list of *Tugenden*:

> *The affirmative aspects*: pride, joy, health, love of the sexes, enmity and war, reverence, beautiful gestures, manners and objects, strong will, the discipline

[2] Here Leiter (2010) may respond by dismissing such passages as rhetoric which 'does not really suggest realism about the content, but rather desperation on the part of the author to reach an increasingly distant and uninterested audience.' However, I would argue that the frequency with which Nietzsche returns to Greek values suggests that he is interested in them not merely for rhetorical purposes.

> of high spirituality, will to power, gratitude toward earth and life – everything that is rich and desires to bestow and that replenishes and gilds and immortalises and deifies life- the whole force of *transfiguring* virtues, everything that declares good and affirms in word and deed.
>
> (KSA 13, 14 [11])

The cultivation of such virtues surely requires an element of judgement. Even if an individual qua 'type' is predisposed to 'deem this to be good,' this very idea of 'deeming' seems to call for rational engagement on the part of the individual. Leiter's notion that valuing, as an expression of 'type-facts' does not occur as an instance of cognising will find the following passage to be problematic:

> The misunderstanding of passion and reason, as if the latter were an independent entity and not rather a system of relations between various passions and desires; and as if every passion did not possess its quantum of reason.
>
> (WP 397)

Leiter's interpretation certainly accounts for the first half of this passage, but neglects to incorporate the second half: 'as if every passion did not possess its quantum of reason'. Here, it would appear that Nietzsche is indeed suggesting that rational activity also has a part to play in action, and that it interacts with the drives and affects. Leiter's account filters out any notion of conscious thought, and reduces it to mere epiphenomenon, but it is not obvious in the above passage that Nietzsche is adopting such a view. Leiter's model of the self therefore seems overly crude: a closer examination of the drives, affects, and conscious thought is required.

Leiter's naturalistic account, by ruling out the idea that conscious thought has any causal efficacy, appears deterministic[3], and as such, does not allow any room for self-formation,

[3] Leiter may object and try to soften his determinism somewhat by adding that an individual's type-facts enumerate a *range* of possible life trajectories, and that external factors such as 'environment and circumstances' can *significantly* accentuate or inhibit the expression of an individual's type-facts. However, no attempt is made to quantify or qualify these external variables, nor is it made clear whether it is even

development, or amelioration. He straightforwardly interprets Nietzsche's formulation 'how one becomes what one is' in *Ecce Homo* as fatalistic, suggesting that this is carried out by 'making no special effort *directed towards that end,* because one becomes what one is *necessarily*' (2002: 86). Not only does this interpretation make the formulation circular and uninformative, i.e. what one is, is whatever one becomes, but such a reading does not quite tie in with the fact that Nietzsche often renders the formulation in the imperative form. Employing the formulation as an injunction suggests that Nietzsche is attempting to persuade us to *act*- to *do* something, and that if we fail to deliver, then there is the chance that we will not 'become what we are.'

The notion of self-formation certainly appears in Nietzsche's writings, for example his comments on figures such as Goethe and Wagner. Nietzsche writes of Goethe:

> He surrounded himself with nothing but closed horizons; he did not sever himself from life, he placed himself within it; nothing could discourage him and he took as much as possible upon himself, above himself, within himself. What he aspired to was *totality*; he strove against the separation of reason, sensuality, feeling, will…he disciplined himself to a whole, he *created* himself.
>
> (TI 'Skirmishes' 49)

Here, there seems to be an emphasis on self-determination and self-formation: active processes requiring some sort of agency on the part of the individual. Again, we should note the inclusion of reason in this process: 'he strove against the separation of reason, sensuality, feeling, will'. Leiter's model thus seems insufficient to incorporate Nietzsche's thoughts on Goethe. The passive, static self which the naturalistic account depicts, and its

---

appropriate to try and calibrate the external factors and internal type-facts in terms of each other. Ultimately the influence of external factors on Leiter's account emerges as being negligible and it seems that type-facts are wholly responsible for what we become. This conclusion is supported by the fact that Leiter substitutes 'trajectories' (note the plural) for the singular 'Fate' which suggests that the trajectory of the individual's life is completely specific: 'the trajectory of a person's life follows a necessary course, one determined by the constitutive type-facts, in interaction with environment and circumstances' (Leiter 1998: 236).

exclusion of conscious thought in agency, is unable to account for the process of self-formation exemplified by individuals such as Goethe.

The Transcendental Account (Gardner 2009)

Gardner (2009) takes the debate regarding the self in Nietzsche to revolve around the issue of normativity. He draws attention to the disunity of theoretical and practical thinking (especially self-formation) in Nietzsche's writings, suggesting that this could only be resolved by invoking a transcendental 'I', before arriving at an ultimately aporetic conclusion. In this way, he offers a very different account to that of Leiter's, and takes seriously Nietzsche's idea of 'becoming what you are', interpreting this as a call for the creation and legislation of values.

Gardner argues that, with regards to theoretical reason, in works such as *Beyond Good and Evil* ('On the Prejudices of the Philosophers'), Nietzsche appears to express naturalistic tendencies. The 'I' is described as an illusion and a fiction, and the self is conceived of in terms of drives and affects, with any consciousness being epiphenomenal:

> I don't concede that the 'I' is what thinks. Instead, I take the *I itself to be a construction of thinking,* of the same rank as 'matter', 'thing', 'substance', 'individual', 'purpose', 'number': in other words to be only a *regulative fiction* with the help of which a kind of constancy and thus 'knowability' is inserted into, *invented into*, a world of becoming…but up to now philosophers have believed, like the 'common people', that in '*I* think' there lay something or other of unmediated certainty and that this 'I' was the given cause of thinking….However habituated and indispensable this fiction may now be, that in no way disproves its having been invented.
>
> (KGW VII- 3.248)

However, Gardner argues that such views become highly problematic when we consider what Nietzsche has to say about practical reason.

Gardner (2009:7) suggests that the kind of subject that realises value (cf. the Sovereign Individual passage in the *Genealogy* or passages throughout *Zarathustra*), must a) realise

himself – his self – as the *ground* of the value that he affirms: 'it is part of what it is to entertain and affirm values in the proper non-alienated, explicitly legislative mode, that one's own contribution, the subject's act of *sponsoring*, be understood as constitutive of the 'object' that comprises one's value' and, b) experience a reciprocal relation between valuing, self-formation, and self-affirmation: 'to determine such and such to be of value is to determine *oneself*, and to affirm oneself *by way of* affirming (p.9) what one values, and vice versa.' Both of these necessary conditions for what it is to be a subject that values, suggest that the 'I' plays a 'fundamental, pervasive, and ineliminable role' for Nietzsche.

The naturalistic picture of the self à la Leiter does not, according to Gardner, allow for either of these conditions to be satisfied. The self qua bundle of drives and affects does not produce first-person thinking in practical or axiological contexts:

> while it is true that there is no inconsistency between *our* thinking of some individual as bearing value on account of their psychological structure, if that individual is to think of *himself* as bearing value, then the I-conception is indispensable: Nietzschean man must set value on *himself*, not on some psychological structure.
>
> (2009:8)

The naturalistic picture is too impersonal: the 'I' qua an epiphenomenal product is a weak concept, and does not seem to serve any purpose. As such, it is totally inadequate for what Gardner regards as Nietzsche's ultimate philosophical purpose, i.e. to forge 'individuals who set value on (affirm) themselves' (2009:8) In order to carry out this project, Gardner maintains that Nietzsche needs an 'I' which has the capacity for self-awareness in judgement: this conception of the 'I' stands in stark contrast to the sub-personal 'I' in the naturalistic account. According to Gardner the drives/affects lack the requisite cognitive and explanatory character needed for valuation.

If we wish to aim for a unity of theoretical and practical thinking in Nietzsche's works, the solution Gardner proposes lies in a conception of the 'I' as transcendental, and he maintains that 'to transcendentalize Nietzsche is not to kantianize him' (2009:19). Gardner is careful to argue that the 'I' which Nietzsche argues against is the specific, hypostatised, platonised conception of the 'I' that plays a role in Christianity, rather than

a transcendental 'I' *per se*, and that Nietzsche's thought contains a 'buried transcendental dimension' (2009:18), which can help us to clarify his views on moral psychology.

Gardner suggests that a conception of the 'I' that Nietzsche needs, must be one which 'holds together in a coherent manner *both* the unitary I of self-consciousness *and* the psychological manifold' (2009:12). Such an 'I' would maintain both: a) that the causality of the 'I' is properly interpreted as necessarily involving the expression of the dominant drive in the psychological composite, and b) that 'the thought, which the Nietzschean subject must entertain when a power-unit realizes itself successfully, wills values, etc., is an *I*-thought, not the thought *that* such and such a power-unit or whatever prevails presently' (2009:12)[4] cf. Korsgaard: 'to regard some movement of my mind or my body as *my action,* I must see it as an expression of my self as a whole, rather than as a product of some force that is at work *on* me or *in* me' (2009:18).

So how does Gardner understand this 'I'?

Gardner firstly delineates the 'I' in negative terms. He understands the 'I' 'not as substratum, nor as equipped with freedom of will in the indeterministic sense, yet as occupying the position of *ground*' (2009:8). Without such an 'I', he suggests that the individual would be left with a 'profound self-alienation', and the normative dimension of valuation would be undermined, i.e. 'the possibility of its being thought that the valuation in any sense 'gets things right'. He does not outline, however, what exactly comprises the Nietzschean 'I will' in relation to values. He suggests that it 'amounts to more than the practical commitment that is implied trivially in any recognition of a norm, and appears to include an aesthetic or quasi-artistic dimension—the Nietzschean subject will relate to his values in something of the way that an artist relates to his works, as something distinct from himself, yet as an object of pride and ownership, for which the subject claims responsibility or with which it identifies' (2009:8).

[4] Here we might compare the Nietzschean subject to the Freudian analysand who would wish to retain first-person thinking over and above the psychological apparatus of the ego, super-ego, Id – there is still a desire to think in I-terms.

After constructing a possible 'I' which he believes could potentially provide a unity of theoretical and practical reasoning in the Nietzschean subject, thus lending credence to Nietzsche's views on normativity, Gardner then notes that Nietzsche does not actually aim to provide a unified theory of self. He then takes this to mean that Nietzsche is illustrating the unavoidable disunity of philosophical reasoning through an aporetic conclusion. In other words, we may wish to reject transcendental realities at a theoretical level, yet crave them or employ them at a practical level. It is not that Nietzsche fails to acknowledge or appreciate the conflict between naturalism and transcendentalism, or to choose between them: Gardner is suggesting that Nietzsche is drawing our attention to the fact that our search for a unity of reason is futile.

But is this really Nietzsche's aim? As Anderson suggests Gardner's 'result seems to be based primarily on an *a priori* argument identifying alleged presuppositions of Nietzschean positions, rather than any direct argument from Nietzsche's texts. As such, it might be thought to tell us more about the shape and force of *Gardner's* post-Kantian commitments than it does about Nietzsche's own view' (2012:4).

I would also add that in Gardner's account, much like Leiter's, the drives and affects are the only components of the self which receive attention, and the level of attention seems somewhat insufficient, resulting in too crude a model of the self. Gardner contrasts Nietzsche's drives and affects with the contemporary anti-realist view of the self as found in Dennett's works:

> In Dennett, for example, the 'I' is regarded as a theoretical posit which is introduced in consequence of taking up the intentional stance, as a kind of conceptual corollary of belief and desire explanation, somewhat in the way that the laws of mechanics warrant the positing of a centre of gravity in physical objects.
>
> (2009:3)

Nietzsche's drives however, according to Gardner's interpretation 'lack at root the cognitive, explanatory character assumed by Dennett'. If this is true, then Gardner is right to wish to posit some kind of 'I' qua conscious capacity over and above the drives to account for the self-awareness needed in Nietzsche's practical thinking, i.e. valuing. But

perhaps we ought to a) take a closer look at the structure of drives and affects, and establish their characteristics, and b) explore the other components of a Nietzschean self. It may be that Nietzsche does indeed provide us with the materials to make sense of his practical philosophy.

The next section will examine two accounts which aim to examine the operations of a Nietzschean self in greater detail: Katsafanas (2012) focuses on the role of consciousness, whilst Janaway (2009) explores the role of the affects.

1.2 Conscious Willing (Katsafanas 2012)

Katsafanas (2012) explores one component of the self in particular, namely conscious willing, and examines whether or not Nietzsche is able to account for this activity. This kind of activity seems important for notions of self-formation. In exploring the idea of conscious willing, Katsafanas offers a richer account of the Nietzschean self, with a particular emphasis on its cognitive capacities.

Katsafanas makes the claims contra Leiter, that conscious thought: a) is not epiphenomenal and, b) has causal efficacy. In his earlier (2005) paper he highlights the importance of consciousness in Nietzsche's thought, when he suggests that, without it, we would be unable to explain the transformation of bad conscience into guilt (as illustrated in *On the Genealogy of Morality*). Katsafanas, in his interpretation of the Nietzschean model of the self, allows conscious thought to have a role in agency, something which Leiter flatly denies, and which Gardner would wish there to be in an account of the self. In order to show this, Katsafanas devotes much attention to the structure of drives, and their relationship to affects and conscious thought, as he attempts to trace the pattern of causality.

Why does Katsafanas insist on demonstrating the possibility of reflective agency in Nietzsche's thought? At this point Katsafanas (2012:6) draws our attention to Nietzsche's Sovereign Individual. The Sovereign Individual is characterised by the capacity 'to promise' which is outlined in GM II: 2. Such a capacity presupposes a certain conception of willing, which Katsafanas goes on to suggest is fundamentally reflective.

In promising, the promisor is able to, as Richardson suggests 'insert a pause' in which to consult his commitments during the process of willing (2009: 139). This pause sees the promisor display a 'strong inhibitive power to refrain from acting immediately upon one's drives'. This notion of a pause or space does seem to surface in Nietzsche's writings:

> [B]etween the original 'I will', 'I shall do this', and the actual discharge of the will, its act, a world of strange new things, circumstances, and even acts of will may be interposed, without causing this long chain of will to break.
>
> (GM II.1)

The Sovereign Individual qua promisor is 'strong enough for that', and is able to preserve his commitments when faced with these temptations. Later in *Twilight of the Idols* Nietzsche identifies willing with the power: 'not to react immediately to a stimulus, but instead to take control of the inhibiting, excluding instincts…the essential thing here is precisely not 'to will', to be able to suspend the decision' (TI 'What the Germans lack' 6).

Katsafanas then suggests that willing here is associated with the capacity to control one's behaviour *reflectively* (2012:7), referring to phrases such as 'gaining control' over instincts, 'suspending decision', and having a 'protracted, independent will' as eliciting images of reflective processes. He also points out that to view Nietzsche's claims about strong and weak wills as referring to mere conflict of desires, does not uncover the full weight of these claims. Katsafanas gives the example of how animals can have desires that are stronger than others, e.g. the desire of self-preservation can be stronger than another such as hunger, meaning that animals with desires 'would *eo ipso* have a strong will' (2012:8). Consequently, when Nietzsche talks of human beings having strong or weak wills, he must be referring to a reflective capacity, not just a conflict of desires.

Katsafanas is then left to explain how we ought to account for the passages where Nietzsche suggests that conscious thoughts, decisions, acts of will play no role in causation of our actions.

> *Error of false causation…*We believed that our acts of will were causally efficacious…Nobody doubted that consciousness was the place to look for all the *antecedentia* of an act, its causes, and that you would be able to find these

> causes there as well - under the rubric of 'motives'…Nowadays we do not believe a word of it.
>
> (TI 'The Four Great Errors' 3)

How are we to resolve this tension? Again, as Katsafanas shows, this can be solved by establishing the specific notions of the will Nietzsche rejects or accepts: we must be careful not to assume that Nietzsche is targeting will *per se* in his writings. In the above passage, he rejects one particular conception of the will (cf. BGE 21 'a will that is ''causa sui'': a will that is determined by nothing other than the agent himself, a will whose ''causes could be found within our own consciousness"), but later in *Twilight of the Idols* he endorses an alternative conception of the will, i.e. a 'strong' but not causally isolated will (TI 'What the Germans lack' 6).

In this way, as Katsafanas notes, it is especially important to take into account any modifications in Nietzsche's views. In early and middle period works up to *Daybreak* Nietzsche appears to be an incompatibilist with regards to free will, maintaining that the notion of a will that is *causa sui* is not possible, instead our actions are causally determined by many factors. He then makes the move to eliminativism, shifting from the claim that we lack *free* will to the claim that we lack will:

> Perhaps there exists neither will nor purposes, and we have only imagined them. The iron hands of necessity which shake the dice box of chance play their game for an infinite length of time; so that there have to be throws which exactly resemble purposiveness and rationality of every degree. Perhaps our actions of will and purpose are nothing but such throws…
>
> (D 130)

However, in the later work *Beyond Good and Evil,* Nietzsche seems to recognise that the quick move from 'our actions are causally determined' to 'our actions are unfree' is not a legitimate one, and adopts a compatibilist position, i.e. our actions could be both causally determined and the products of our wills. In BGE 21, he refutes the idea of a will that is *causa sui*, but then does not move on to eliminativism, as he did in the earlier works:

> Suppose someone were thus to see through the boorish simplicity of this celebrated concept of ''free will'' [as incompatibilist freedom] and put it out of

> his head altogether, I beg of him to carry his ''enlightenment'' a step further, and also put out of his head the contrary of this monstrous conception of ''free will'': I mean ''unfree will,'' which amounts to a misuse of cause and effect…The ''unfree will'' is mythology; in real life it is only a matter of strong and weak wills.
>
> (BGE 21)

So how should we understand the role of conscious thought in a Nietzschean self?

Before sketching out exactly how conscious thought operates in Nietzsche, Katsafanas is careful to specify that the model of agency which Nietzsche attacks is Kantian[5] or Lockean. In such models, self-conscious reflection can allow deliberate suspension of motives in that it enables us to distance ourselves from our motives, ensuring that these motives cease to 'dominate' us. Once the motives have been 'suspended', we may 'consider' and 'weigh' them, and deliberate as to whether they are good reasons to act. Nietzsche, as we have seen in passages mentioned (BGE 3), is extremely sceptical of reflective agency: an individual who acts reflectively is very much still 'secretly guided and channelled by his drives', and that in reflective speculation, an agent's intellect is merely 'the blind instrument of another drive' (D109).

In his 2005 article 'Nietzsche's Philosophical Psychology', Katsafanas closely examines the structure of drives. He suggests that drives have a crucial evaluative component, and defines a drive as a 'disposition that manifests itself by informing an agent's perception of a given object, generating an evaluative orientation towards that object and thereby bringing about action.' In his 2012 article 'Nietzsche and Kant on Reflective Agency', Katsafanas discusses how both drives and conscious thought play a role in agency. If we return to this passage:

> The misunderstanding of passion and reason, as if the latter were an independent entity and not rather a system of relations between various

[5] Katsafanas is here replying to Allison's (1983) and Barron's (2012) interpretations of Kant.

> passions and desires; and as if every passion did not possess its quantum of reason.
>
> (WP 397)

We can see how Katsafanas, unlike Leiter, attends to both halves of this claim, i.e. that reason is not independent from drives, and that drives are not independent from reason: 'the passions as influenced by reflection: the passions do not enjoy any independence from reflection; on the contrary, they are everywhere influenced by reflection' (2012:25). He suggests that for Nietzsche, reflection: a) does not operate in an instantaneous manner (as we so often believe it to do), but rather its effects are gradual and incremental and, b) is just one causal factor amongst others.

In order to further elucidate these claims, Katsafanas first offers a breakdown of the Kantian model of willing, which he believes can be arranged into the following three stages:

1. (Suspension) When an agent reflects on her motives for A-ing, she suspends the influence of the motives upon which she is reflecting, i.e. the agent is able to still these motives and choose in independence of them.

2. (Inclination) In deliberative agency, motives incline without necessitating. The agent's motives could be the same, and yet she could choose differently, i.e. given Suspension the agent is able to reflect on motives and suspend the motivational aspects of these motives; consequently, motives do not necessitate any actions.

3. (Choice) Typically, if I am faced with two actions that it is possible for me to perform, A-ing and B-ing, and I choose to A, then I will A, i.e. the agent takes her choice to determine what she will do.

Katsafanas then proposes that Nietzsche rejects certain components (Suspension) of the above model of willing, but retains others (Inclination and Choice):

> Due to the fact that he develops a more complex account of motivation, Nietzsche concludes that reflection is not capable of suspending the influence of motives. Nonetheless, he agrees with Kant that motives do not determine choice: our motives could be the same, and yet we could choose differently.

> Moreover, he maintains that conscious choice plays a causal role in the production of action.
>
> (2012:3)

Suspension does not hold for Nietzsche due to the fact that our actions are driven by underlying motives, of which we are not always aware. Moreover, even if we find ourselves identifying a particular motive, and subsequently 'suspending' it, all that has happened is that the motive has in fact become more covert 'operating through reflection itself.' In BGE 117 Nietzsche writes that 'the will to overcome an affect is ultimately only the will of another, or several other, affects.'

Katsafanas uses the example of jealousy to illustrate this point:

> The attitude influences the agent's reflective thought itself: the agent experiences herself as having a reflective distance from the attitude, as scrutinizing the attitude and asking herself whether there is a reason to act on it; but, all the while, the attitude influences the agent's reflective thought in ways that she does not grasp.
>
> (2012:13)

Here, the way in which we try to reflect on our jealousy is in itself structured and coloured by that jealousy unbeknownst to us: 'a jealous agent's deliberative process itself can be influenced by these attitudes; they can incline him to draw conclusions that are not supported by the evidence, to give excessive weight to certain features, and so on.' Katsafanas cites Othello as an apt literary example of this process. In this way, a motive like jealousy influences the agent's perception of reasons and indirectly moves the agent. The idea that we can stand apart from our motives therefore does not hold for Nietzsche.

Katsafanas then suggests that Nietzsche is able to reject Suspension yet uphold Inclination and Choice. Having rejected Suspension we may be inclined to think that conscious thought has no role whatsoever in action, but in order to maintain both Inclination and Choice, conscious thought for Nietzsche cannot, *contra* Leiter's interpretation, be causally inert. How are we to understand the causal role of consciousness?

Leiter's claim that consciousness is causally inert relies on two models of the will: 1) 'The Will as Epiphenomenal', whereby conscious mental states and any events associated with willing do not play any part in the causal chain leading to action. This is because on this model both conscious experiences of willing and action are caused solely by the drives and affects. There is no causal connection whatsoever between conscious experiences of willing and action, and 2) 'The Will as Secondary Cause', whereby conscious mental states and events associated with willing are permitted to be part of the causal chain leading to action, but they are not primary causes. Rather, they are only efficacious in virtue of other causes. On this model, the drives and affects cause conscious experiences of willing, which then cause actions.

Katsafanas, however, proposes that these models by no means provide an exhaustive understand of the causal role of consciousness, and that Nietzsche in fact rejects both of these particular models. Instead, he suggests, Nietzsche has a far more complex picture of the relationship between conscious thoughts and affects, in so far as he argues that conscious thought can transform the motivational tendencies of the affects. Such an interpretation attributes a vital role to conscious thought in the production of action.

So how exactly does conscious thought transform the motivational tendencies of the affects?

Katsafanas uses the example of pain and how, on Nietzsche's account, it does not have a determinate motivational impact on agency. This view is made clear in the *Genealogy* when Nietzsche declares that 'what really arouses indignation against suffering is not suffering as such but the meaninglessness of suffering' (GM II.7). Here, it seems that we only experience real aversion to suffering when it is coupled with an interpretation, i.e. 'this suffering is meaningless'. Conversely, we may have a pleasurable experience of pain if it is coupled with an interpretation such as 'this pain means that I am exercising vigorously.' In *The Gay Science* Nietzsche expresses this very thought: 'that a violent stimulus is experienced as pleasure and pain is a matter of the interpreting intellect, which, to be sure, generally works without our being conscious of it; and one and the same stimulus can be interpreted as pleasure or pain' (GS 127).

This notion of conscious thought transforming affects via interpretation does not just apply to the example of pain, but to affects generally. In *Daybreak* Nietzsche writes:

> Drives transformed by moral judgments.—The same drive evolves into the painful feeling of cowardice under the impress of the reproach custom has imposed upon this drive: or into the pleasant feeling of humility if it happens that a custom such as the Christian has taken it to its heart and labeled it good. That is to say, it is attended by either a good or a bad conscience! In itself it has, like every drive, neither this moral character nor any moral character at all, not even a determinate accompanying sensation of pleasure or displeasure: it acquires all this as a second nature only when it enters into relations with drives already baptized good or evil, or is noted as a property of beings that have already been morally ascertained and assessed by the people.—
>
> (D 38)

Here, Katsafanas suggests that Nietzsche views 'motives as causally indeterminate: motives have causal tendencies, but the particular behaviours that they characteristically cause are dependent on the associated interpretation' (2012:20). These 'interpretations' can be understood as involving conscious thought. Moreover, as Katsafanas notes, because one and the same motive, e.g. pity, can be interpreted by two different agents in radically different ways (e.g. one agent via Interpretation A experiences the pity as attractive, and is driven to help another individual in distress, whilst the other agent via Interpretation B experiences the pity as unattractive and ignores said individual), we must also reject Leiter's other model of 'The Will as Secondary Cause' as rendering the chain of causality from motive to action as too simplistic.

If we accept the ideal that conscious thought does indeed play a role in human action, we must then ask in what way does this transformation of affect through conscious thought take place?

There are two crucial aspects of Nietzsche's conception of willing according to Katsafanas. Firstly, the effects of conscious reflective thought are gradual and incremental. Nietzsche observes the transformation of motives via conscious thought across generations: they must be 'constantly internalized, drilled, translated into flesh and reality' (GS 301). This is not to say that this process occurs with ease –Nietzsche

makes it clear that the transformation of motives can be an arduous process 'we have to learn to think differently—in order at last, perhaps very late on, to attain even more: to feel differently'' (D 103). Indeed, at times, he makes it clear that there will be some affects which cannot be transformed: 'learning changes us . . . but at the bottom of us, really 'deep down', there is of course something unteachable, some granite of spiritual *fatum*…' (BGE 231).

Secondly, reflective thought is but one causal factor amongst many others. In order to explain this, Katsafanas distinguishes between what he calls a 'Trigger' model of willing, and a 'Vector' model of willing. The Trigger mode, according to Katsafanas, is what we see in Kant. Here, the agent has various motivational desires and affects which may incline the agent to certain actions, but, after deliberation, these motivational states are incapable of causing the agent to act. It is only after the will has consented to a certain course of action, and endorsed a certain motivational state, that these desires and affects then become causally efficacious. Thus, choice alone determines action and the will acts as a trigger. The Vector model, by contrast, reduces the causal importance of the will: the will is merely one source of motivation amongst others. The will may reinforce other motives 'by placing its motivational weight behind them' but it is not uniquely efficacious. Action is determined by the vector of motives (which includes the will). On this model, motives, episodes of decision, and conscious thought are all seen as causal forces which interact with each other: 'none enjoys a privileged position in the production of action' (2012:28).

This vector model allows conscious willing to have a role in agency (and to retain Inclination and Choice), but does not require that it take the form of some substantive faculty capable of Suspension which we see in the Kantian model. Instead, it is seen to have a rather more modest role than a Kantian conception of the will. Willing is one causal factor amongst others, and its effects are not immediate: instead we see gradual, incremental change (and the transformation of motives is far from guaranteed). The vector model also allows us to make sense of Nietzsche's claims that there are strong and weak wills since conscious thought, though causally efficacious, is not absolutely guaranteed to have a decisive causal impact. Katsafanas notes how the Sovereign Individual would be vastly more capable of controlling his behaviours through conscious

willing: 'the capacity deliberatively to form intentions, and to remain resolute in their realization, is something that might well vary across individuals. Indeed, we already know that in certain cases it does vary: certain individuals seem to manifest more self-control than others' (2012:27).

Katsafanas' account pays close attention to the workings of the psychological apparatus Nietzsche mentions: he offers careful discussion of both the drives and affects, and does not neglect the importance Nietzsche does accord to conscious reflection or willing. As Stern (2015) suggests:

> while instinct *can* indicate something non-conscious, some of these examples of drives and instincts appear to *demand* or *presuppose* something like conscious reflection: mulling things over, intellectuality, researching, reflecting. And one can have a *Hang* for dialectics or for the *vita contemplativa*. Affects, taking into account what Nietzsche tells us, obviously stretch well beyond brief, exhausting 'seizures of excitement' or even 'ways in which we feel' (Janaway 2009: 52) to incorporate desires and demands; indeed, some affects have 'wills' of their own (BGE 117).
>
> (2015:127)

It thus seems vital to incorporate cognitive aspects and the idea of conscious willing into an account of the Nietzschean self.

Perspectival Knowing (Janaway 2009)

Like Katsafanas, Janaway emphasises cognitive aspects of the self in Nietzsche, and focuses, in particular, on the affects, which he takes to the sole locus of interpretations. Janaway takes perspectival knowing to be the main motivation for a 'relatively unified and autonomous self' (2009:1). Self-formation can therefore be seen to include this notion of 'perspectival knowing' on this account. As such, Janaway argues that a Nietzschean self is a self which possesses self-awareness, and a certain level of creative agency. This creative agency consists in controlling the affects and 'shifting them in and

out', something which is required for perspectival knowing, and which implies that Nietzsche does have a reasonably unified and autonomous self in mind.

Before examining these claims regarding the self, we should first examine the idea of perspectival knowing, and how Janaway proceeds from Nietzsche's arguments regarding cognition in order to argue for a self.

In the *Genealogy* , Nietzsche espouses this notion of perspectival knowing:

> there is *only* a perspectival seeing, *only* a perspectival 'knowing'; and *the more* affects we allow to speak about a matter, *the more* eyes, different eyes, we know how to bring to bear on one and the same matter, that much more complete will our 'concept' of this matter, our 'objectivity' be.
>
> (GM III:12)

In this way, Nietzsche rejects prior conceptions of knowledge as requiring the absence of affects or passions, and instead argues not only for: 1) for the inseparability of knowledge and the affects, but also 2) that the greater the number of affects involved in interpretation, the more complete our knowledge will be. Janaway takes Nietzsche to suggest that there is an intrinsic link between affects and knowledge, which can be beneficial, and we should not seek to separate the two: 'to eliminate the will altogether, to disconnect the affects would be to disable knowledge' (2009:3).

These affects possess a cognitive potency in that 'our belief that our feeling shocked, embarrassed, disgusted or attracted by some phenomenon *tells us something about* that phenomenon.' In the *Genealogy,* for example, by examining the role of cruelty in the origin of guilt and bad conscience 'we acknowledge, beneath our more obvious feelings of anger and disgust, a streak of joyfulness in seeing and making suffer' (2009:4). In this way, acknowledging our affects allows us to acquire knowledge about ourselves, and sheds light on our moral attitudes.

That the notion of perspectival knowing can, in this way, allow for greater self-awareness, is for Janaway, crucial for Nietzsche's idea of revaluation and consequently, self-formation. Janaway suggests that the process of revaluation itself is cognitive:

> One *comes to believe* a certain explanation as true, one *judges* a set of psychological states as unhealthy, one *tries* to feel a new set of affects, and *identifies oneself with* specific critical second-order attitudes regarding certain of one's feelings.
>
> (2009:18)

In such a process, it is necessary that the individual possesses some notion of himself as 'deciding, choosing, and trying as a genuine agent'. This is not to say that the agent experiences an 'unlimited possibility of action unconstrained by character and the causal order'. Nietzsche quite clearly does not forward such a conception of free will. However, neither does he advance the notion of total unfreedom of the will. In this way Janaway suggests that Nietzsche wishes to preserve a sense of creative agency 'because without it his proposed critique of moral values and his project of learning to think and feel in healthier ways would make little sense' (2012:18).

So how do the affects lead us to posit a self? It seems that the answer to this question will shed some light on how Janaway envisages creative agency in Nietzsche. In order to answer the question fully we must examine the following in Janaway's account: (i) what is an affect? (ii) what is a drive? (iii) how are affects and drives related? And (iv) does the interplay between drives and affects constitute knowledge?

Janaway defines affects as: 'inclinations and aversions' 'pro and contra' 'for and against' – inclinations or aversions of some sort… affects are, at the very least, ways in which we *feel*'. He also highlights the fact that some affects are unconscious or beyond accurate apprehension. Here he cites D 115 to demonstrate the range of apprehensibility of affects:

> [W]ords really exist only for *superlative* degrees of these [inner] processes and drives. […] Anger, hatred, love, pity, desire, knowledge, joy, pain—all are names for *extreme* states: **(p.53)** the milder, middle degrees, not to speak of the lower degrees which are continually in play, elude us, and yet it is they which weave the web of our character and destiny'.

Janaway, like Katsafanas, accords motivational capacities to affects suggesting that affects play a vital role in interpretation:

> It is our needs which interpret the world: our drives and their for and against. Every drive is a kind of lust for domination, each has its perspective'; and 'Who interprets?—Our affects.
>
> (2009:7)

Drives, meanwhile, are, for Janaway 'relatively enduring tendencies'. He seems to agree with Katsafanas that drives can be characterised as dispositions, but he disagrees with Katsafanas' claim that drives provide agents with structuring goals of which they are ignorant, providing the example of the drive to artistic self-expression, i.e. it need not be the case here that the agent must remain ignorant of the ultimate goal of this drive.

Moreover, it is not the case that this drive must persist constantly. Janaway claims that Nietzsche's idea that man has 'bred' the plenitude of drives into himself implies that drives may be culturally acquired or lost[6]:

> according to Nietzsche's use of 'instinct' and 'drive', such things need not be built unchangeably into human beings—neither generically into humans qua humans, nor into the constitution of any human being considered individually. How an individual's drives operate over time, and even what drives an individual continues to have, is open to change.
>
> (2012:7)

In this way, it seems that Janaway's interpretation leaves us with a more plastic notion of the self than Katsafanas'. It is not just that drives remain latent until activated, waxing and waning in response to conditions, but they can in fact disappear altogether. This plasticity seems important for self-formation particularly in light of passages such as the following from *The Gay Science*:

[6] Janaway's conception of drives places them in a reciprocal relationship with something like morality – an external condition like morality may influence which drives are acquired or lost, which in turn influence the kind of moralities an individual will adhere to or create. In this way Janaway's account, unlike Leiter's, offers a richer explanation of the influence between internal and external factors and is able to account for development of drives/moralities.

> It is practiced by those who survey all the strengths and weaknesses of their nature and then fit them into an artistic plan until every one of them appears as art and reason and even weaknesses delight the eye. Here a large mass of second nature has been added; there a piece of original nature has been removed – both times through long practice and daily work at it. Here the ugly that could not be removed is concealed; there it has been reinterpreted and made sublime. Much that is vague and resisted shaping has been saved and exploited for distant views
>
> (GS 290)

The ability to acquire or lose drives is particularly pertinent here. So, how do affects relate to drives? Janaway suggests that 'we might hypothesize that a drive is a relatively stable tendency to active behaviour of some kind, while an affect, put very roughly, is what it feels like when a drive is active inside oneself' (2009:6). Affects here then seem to be emotional responses to the operation of the drives, and to the successes or failures of drives in obtaining their goals. The constitution of an individual's drives would therefore seem to have a direct influence upon their affective landscape.

We have seen that Janaway believes that the affects play a vital role in perspectival knowing, but we must now ask how can we move from an 'interpretation' to a 'knowing'? On the above definition of the relationship between drives and affects, knowing must equal some activity of the drives which compose the individual, along with the feelings essentially involved in their activity.

Are the drives and affects enough to warrant knowing? Janaway suggests that this is an insufficient model to constitute knowledge. Instead we require operations to take place upon the affects and interpretations, namely the 'shifting them in and out', and 'having them in one's power', which Nietzsche speaks of. Differences in perspectives and affective interpretations must be made clear, and manipulated in some way so that they are useful for knowledge.

Like Gardner, Janaway here raises the problem of the 'I.' He maintains that such operations must be conducted by some kind of self-conscious unity, in order for us to have a coherent notion of controlling affects, and bringing them to bear on one another, and the same subject-matter:

> I have to be aware that affects *A* and *B*, each of which may 'speak' interchangeably about the same subject-matter, are both *mine*. I have to be that which feels both affects, regards itself as feeling them both, and takes some attitude towards its subject-matter in the light now of this affect, and now of that.
>
> (2009:8)

It seems that Janaway, like Gardner, desires some sort of 'I' over and above the drives and affects, which is able to self-consciously adopt attitudes:

> but I have to be, in my own self-conception, a sufficiently unified self that *I* can 'take sides' between the various drives that (though I did not originally will them) I find within myself. Likewise, it is not just that each of the affects I find within myself has a goal of its own, but rather that *I* have a goal in pursuit of which I can flexibly use the affects I feel.
>
> (2009:9)

In this way, the naturalistic model of the self offered by Leiter looks inadequate. So what kind of 'I' does Janaway posit?

The 'I' for Janaway must be a cogniser. Such a conception of the self enables Janaway to shed light on Nietzsche's idea of self-affirmation, and notions like Amor Fati, both important for a project of self-formation. Janaway suggests that a self qua cogniser, who is able to self-consciously adopt attitudes, and possess some degree of self-knowledge, is able to partake in first-order and second-order willing (the latter being crucial for something like Amor Fati).

Janaway (2012) distinguishes between first-order and second-order willing, stating that animals have only the former, but it is the fact that humans have the latter which enables them to take up attitudes of self-affirmation. He concedes that 'no human being has complete knowledge of his drives, and no one is in full rational control of their presence or mode of expression' (2012:11), but that we may still take up attitudes to ourselves. Amor Fati requires that an individual would be well disposed to desire their life again, embracing all aspects of it, including all imperfections. Whilst the idea of desiring the undesirable may be difficult to grasp, through the idea of second-order willing we can

begin to see how this process might be feasible: 'numerous events in any life will be undergone, remembered, or anticipated with a negative first-order attitude; but that is compatible with a second-order attitude of acceptance, affirmation, or positive evaluation towards one's having had these negative experiences' (2007:257-8).

In order to carry out this second order willing, Janaway argues that a self qua cogniser, which is able to have the affects within his control and manipulate them 'in accordance with a higher goal' of the revaluative project, is necessary. In this way, revaluation and self-affirmation depend on a relatively unified and autonomous self that can stand over and above the affects. Janaway is careful to note that this is a self which is not to be achieved by everyone (and not even to be achieved to a full extent), only the Sovereign Individuals:

> The fact that self-knowledge (likewise selfhood as such, as Gemes shows 36) is hard, that most have very little of it, and even that no one ever attains it completely—none of this shows that self-knowledge is impossible: only that it is rare among human beings, that it is a task set for a few of us rather than a given, and that its achievement is a matter of degree.
>
> (2009:17)

Janaway aligns this kind of cognitive achievement with the notion of unification - *Einheit.* This idea of a unified multiplicity is emphasised by Nietzsche in several works. In *Beyond Good and Evil* he states that 'only this should be called greatness: the ability to be just as multiple as whole, just as wide as full' (BGE 212), and in *Twilight of the Idols* he cites Goethe as the ideal unified individual: 'what he wanted was *totality*; he fought against the separation of reason, sensibility, feeling, will…he disciplined himself to wholeness' (TI 'Skirmishes', 49).

Nietzsche frequently considers greatness in terms of the properties of and relations between an individual's drives and affects. He discusses: a) the strength of drives, b) their plurality, and c) the fact that they are held in a unity despite being in constant conflict with one another, e.g. 'the highest man would have the greatest multiplicity of drives, in the relatively greatest strength that can be endured. Indeed, where the plant "man" shows himself strongest one finds instincts that conflict powerfully (e.g. in Shakespeare), but are

controlled' (WP 966 cf. BGE 212, TI 'Skirmishes' 49). Such a model of unity is reminiscent of Heraclitus' *diapheromenxon-sympheromenon* ('internal variation' and 'coherence').[7] In this way the self appears as a scaled down Heraclitean cosmos.

The notion of improving and cultivating this *Einheit,* thus achieving the 'task' of the self, can be explained by Janaway's account of self-affirmation through second-order willing: 'in the human case there is the possibility of attaining a greater degree of unity in the process of taking attitudes to oneself' (2012: 11). He uses the example of an individual with a strong sexual drive or a strong drive to artistic self-expression:

> While these drives persist, the agent might be continually striving to *disown* them, having set him- or herself to be abstinent and socially conforming. Might not the human being who willed themselves to Goethean wholeness be someone in whom, by contrast, such conscious striving against drives was absent, and whose will aligned itself with as many of the drives as possible, thereby constituting the drives as more of a unity?
>
> (2012:11)

How does this alignment take place? Through the manipulation of affects via the self qua cogniser. Janaway thus takes self-formation to have a fundamentally cognitive element with the ultimate aim of unifying a multiplicity of perspectives and attitudes.

The next section will further explore the notion that the self is a task or achievement. Whilst this kind of account is opposed to the fatalism of Leiter's reading, we will see that this achievement is only possible for the few.

[7] Heraclitus Fragment 10 'Couples are things whole and not whole, what is drawn together and what is drawn asunder, the harmonious and discordant. The one is made up of all things, and all things issue from the one.'

1.3 Autonomy and Agency (Gemes 2009)

Similarly to Janaway, Gemes (2009a) places great emphasis on the concept of *Einheit*, and takes this unity and personhood to be an achievement or task. Gemes puts forward a naturalistic-aestheticist model of the self. Like Leiter, Gemes holds that consciousness in Nietzsche is merely epiphenomenal (meaning that, unlike Katsafanas and Janaway, there is little talk of conscious thought playing any motivational roles in this account), and that there is no 'I' above and beyond the structure of the drives and affects: 'this for Nietzsche would be an empty mere geometric point' (2009a:17)

However, this is not to say that we must dispense with ideas of self-determination and self-awareness. Gemes regards Nietzsche's comments on free will to be especially pertinent to our ideas of the self, and takes the debate about the self to be motivated by claims about agency: what kind of self is capable of genuine agency? How should we understand agency? In answering these questions, Gemes differentiates between two accounts of free will: 1) deserts free will (rejected by Nietzsche) and, 2) agency free will (accepted by Nietzsche). This distinction proves to be helpful in explaining apparent inconsistencies in Nietzsche's writings as regards free will: 'the denials are denials of deserts free will and the invocations are invocations of agency free will' (2009a:2).

Gemes points out that whilst Nietzsche totally rejects the notion of free will *causa sui,* he equally rejects the idea of an unfree will:

> Suppose someone were thus to see through the boorish simplicity of this celebrated concept of "free will" and put it out of his head altogether, I beg of him to carry his "enlightenment" a step further, and also put out of his head the contrary of this monstrous conception of "free will": I mean "unfree will," which amounts to a misuse of cause and effect.
>
> (BGE 21)

Gemes suggests that Nietzsche's rejection of the traditional free will is due to the fact that it usually works in such a way as to instil guilt, resulting in a passive attitude to the world: 'the rule of the 'Thou shall not's''. He also sees an unfree will as inimical to active engagement with the world, as it leads to fatalistic pessimism: 'a fatalism of the weak

willed.' Such statements from Nietzsche seem to be problematic for Leiter's fatalistic interpretation- it seems that Leiter is operating in the free/unfree wills paradigm, rather than the strong/weak wills paradigm which Nietzsche calls for.

So how should we understand agency given that there are strong/weak wills? And how does the resultant picture of agency relate to the self? Given that Gemes thinks the self or personhood is something to be achieved, this implies that we must have some degree of freedom to carry out the task. On Gemes' account freedom equates to self-determination. So how can we determine ourselves?

Gemes distinguishes between deserts free will and agency free will. Deserts free will relates to the question of whether an individual could have done otherwise, and is tied to ideas of punishment and reward. This is rejected by Nietzsche for the reason that he does not accept a free will *causa sui*. In *Human, all too Human* Nietzsche writes of the 'fable of intelligible freedom':

> Now one finally discovers that this nature [of man] cannot be responsible, since it is completely a necessary consequence and is assembled from the elements and influences of past and present things; consequently, one is not responsible for anything, not for his nature, nor his motives, nor his actions, nor his actions nor for his effects. Thereby one achieves the knowledge that the history of moral sensations is the history of an error, the error of responsibility which rests on the error of freedom of the will.
>
> (HAH 39)

It is thus clear why Nietzsche would not accept the notion of deserts free will, since agents cannot be responsible in the way that deserts free will would demand, i.e. the agents do not possess a will autonomous from the causal order. The idea of punishment and reward thus does not make sense. Later Nietzsche writes that:

> for he who is punished does not deserve punishment; he is merely being employed as the means of henceforth deterring others from certain action; likewise, he who is being **(p.36)** rewarded does not deserve his reward: for he could not have acted otherwise than he did.
>
> (HAH 105)

We may compare this with the analogy of the lambs and birds of prey in the *Genealogy* (GM I:13) - neither can act otherwise.

Nevertheless, Gemes argues that Nietzsche 'wants to reject the notion that in doing such and such one might have done otherwise, yet he wants to affirm that genuine agency is possible, if only for a select few' (2009a:2). Agency free will may therefore be more appropriate. This conception of free will relates to the question of agency itself, i.e. what constitutes an action as opposed to a mere doing. Gemes suggests that 'the type of freedom Nietzsche is invoking here does not involve freedom from the causal order, nor is it bound to questions of deserts. Plainly it is tied to the question of what is it to have genuine agency' (2009a:7). Passages relating to the Sovereign Individual clearly mention autonomy in some sense; in the *Genealogy* Nietzsche outlines the Sovereign Individual as one who is 'autonomous… the man who has his own independent, protracted will…the master of a free will' (GM II:2).

So, how should we understand this autonomy?

Gemes suggest that there is a parallel between Nietzsche's notion of free will and Hume's claim that a 'liberty of spontaneity' (free will) is consistent with a denial of 'liberty of indifference' (determinism). Hume claims that a free act is one which stems from one's character or dispositions. Gemes claims that the same thought occurs in Nietzsche albeit with a more robust account of character as 'a stable, unified, integrated, hierarchy of drives' (2009a:6): The freest act is that in which our own most strongest, most finely practiced nature springs forth, and in such a way that at the same time our intellect shows its directing hand. (KSA 10:258, Spring–Summer 1883)[8].

Gemes then goes on to say: 1) this 'freedom' is a very demanding condition that most of us will fail to meet since not all of us are genuine selves i.e. Sovereign Individuals cf. '"people" . . . the "person" is a relatively isolated fact (KSA 12:491), and 2) genuine agency

[8] Noticeably Gemes does not discuss the role of 'intellect' in his account. It is not clear whether he assumes that Nietzsche takes it to be a separate faculty above or on a par with the drives/a drive itself and there is no discussion of how we ought to understand the relation between intellect and nature. In this quotation, it looks as if Nietzsche is separating intellect and nature, and moreover assigns intellect a directive role.

and freedom is only possible under the right conditions. So, on this account freedom and agency are open only to the special few and even then, the right conditions must persist in order for this to be achieved.

Gemes does not specify exactly what these conditions are but it is made clear that such 'fortuitous circumstances' (2009a:11) involve a combination of the internal (a strong master drive), and the external. Should these conditions obtain then those with a 'strong will' will exercise genuine agency and self-determine through taking command of, ordering, and organising lesser drives. This is what Gemes understands Nietzsche to take genuine selves to be: 'to indulge the fable of "unity", "soul", "person", this we have forbidden: with such hypotheses one only covers up the problem' (KSA 11:577, June–July 1885). The problem which this being covered up is, according to Gemes, that 'unity, soul, personhood are not pre-given existences but rare achievements to be gained by hard effort. The fable is in the notion that these have already been achieved, that we as mere humans already have unity, are persons' (2009a:13).

Presenting agency, freedom, and personhood as tasks to be achieved is according to Gemes, less of a serious metaphysical claim than a means of engendering 'active creative engagement' with the world: 'Nietzsche aims to change his preferred readers from being mere conduit points of a vast array of conflicting inherited drives into genuinely unified beings' (2009a:14).

Gemes' account seems to focus more on Nietzsche's ideas on freedom and agency, rather than outlining a detailed model of the self. Several elements seem to be somewhat neglected namely: 1) the distinction between drives and affects, and the relationship between them, 2) the notion of first-person subjectivity, and 3) the role of consciousness, meaning that we are left with a rather thin, impersonal account of the self.

Gemes does not make any claims to be discussing a first-person account: 'first, it does not seek to explain the phenomenology of free will, what if feels like when one feels one is freely choosing to do A rather than B. Second, it does not place the subjective I at the core of the account' (2009a:16). This is because he views character as being glossed in terms of the ordered multiplicity of drives. Consciousness is reduced to an epiphenomenon on Gemes' account: 'Nietzsche is on the whole fairly dismissive of the import of

consciousness' (2009a:16), and he refers to this passage in *The Gay Science* to support this claim:

> Leibniz's incomparable insight that has been vindicated not only against Descartes but against everybody who had philosophized before him—that consciousness is merely an accident of experience and not its necessary and essential attribute.
>
> (GS 305)

However, it seems that like Leiter, Gemes does not address many other passages which suggest that consciousness cannot so easily be reduced to epiphenomenon.

Gemes' account sees Nietzsche as outlining very narrow specifications for what it means to be a self. It appears that only the Sovereign Individual can be classified as a self, and as capable of achieving the task of self-hood. In this way, self-formation emerges as a very exclusive phenomenon on this account.

Repository Self (Anderson 2012)

Like Gemes, Anderson (2012) also sees the self in Nietzsche as something to be carried out and achieved, and the idea of self-formation as self-governance is central to Anderson's account. However, unlike Gemes, Anderson takes the Nietzschean self to be a distinct psychological object, conceived of as an 'I' over and above the drives. In his account, he argues that Nietzsche outlines a thicker notion of the self than is seen in naturalistic interpretations- Anderson terms it a 'minimal self.'

This minimal self is seen as a repository of drives and affects, and its boundaries are porous. Like Janaway, Anderson believes the self to possess a degree of plasticity in that 'there is nothing to prevent my forming and acquiring new drives and affects, nor driving some of the ones I have out of existence' (2012:17). In this way, Anderson wishes to account for the influence of external factors upon the drives and affects. Anderson takes the debate to be motivated by concerns regarding autonomy in Nietzsche:

> Autonomy is central to the rare form of strong individuality he praises: the free spirit he idealizes is supposed to be independent from custom and tradition; she 'creates herself' precisely in the sense of giving herself values or laws of her own; she has 'independence of the soul' (see GS 98, 99, 335, 347; BGE 29, 41, 43–4, 203; GM II, 1–3; *et passim*).
>
> (2012:3)

Autonomy is thus tied to self-governance and self-formation.

How exactly should we understand the 'I over and above the drives' in Anderson's account? He specifies that the self as 'a whole is something over and above the constituent drives and affects, but it is not a simple, essentially unified and conscious, transcendental ego, which is fundamentally different in kind from the attitudes that compose it' (2012:17). In this way, Anderson resists both naturalistic and transcendentalist readings.

Unlike Gardner, Anderson does not see the need for a transcendental self in Nietzsche's account: 'this minimal self lacks the strong features of a transcendental 'I'; it is complex, not simple, and its boundaries do not coincide with those of consciousness' (2012:1). By envisaging the soul as a multiplicity 'soul as society of drives and affect' (BGE 12) Nietzsche denies that 'whatever is subjective at all must exhibit a strong and essential unity proper to consciousness as such, and thus to deny that there is any need to postulate a unified transcendental ego' (2012:8). Gardner's claim that there is a 'buried transcendental dimension' in Nietzsche's thought thus appears problematic.

Transcendentalist accounts like Gardner's are motivated by normative concerns regarding value creation. For an individual to create values, a precondition is for them to have a conception of themselves as a unified practical agent who is the source of these values: 'even if the values she posits are influenced by the drives within her, the individual self must (first-personally) think of them as her own—and not merely the demands of some dominating drive—on pain of a 'profound self-alienation' (Gardner 2009:9)' (2012:4). Here the idea of possession of drives in a personal manner is important.

However, it is not obvious that this is the only way of thinking about preconditions for value creation. Like Katsafanas, Anderson suspects that the psychological apparatus Nietzsche provides, actually contains the resources for these preconditions for the value

creation which is essential to self-formation. He suggests that the explanatory resources of Nietzsche's moral psychology are infinitely richer and more diverse than a naturalist belief/desire psychology, and it would seem strange to ignore the complexity of the psychological apparatus Nietzsche offers.

How does Anderson understand the drives and affects?

Here, Anderson argues that in terming the soul a 'social structure of the drives and affects' and a 'subjective multiplicity', Nietzsche resists strong eliminativism or any reductionist positions (contra naturalistic interpretation), about the relation between the self and its constituents. According to Anderson the mentions of 'social structure' implies:

> something that goes beyond the individuals who participate in it – a more or less definite group reality that may or may not characterise those individuals and their relations' and 'subjective multiplicity' suggest 'a structure with the subjective capacity to inhabit attitudes of its own, including, potentially, attitudes towards its constituent drives and affects.
>
> (2012:10)

These characterisations of the soul heavily imply for Anderson that the self has some emergent reality over and above its constituent drives and affects.

Like Katsafanas, Anderson views drives as having two-place complements (in contrast to something like perception of an object which has a one-place complement), and extends this model to affects, i.e. drives have an aim/object structure and affects have a stimulus object/default behavioural response: 'the attitude itself colours the salience and evaluation of the stimulus object and it governs both the pattern and the manner of the agent's default response' (2012:12). Both drives and affects already come 'evaluatively pre-loaded', i.e. 'the feeling component of affect carries evaluative baggage **(p.220)** that shapes and colours our perception of the stimulus and governs the manner characteristic of the default action path it suggests to us' (2012:13).

Anderson then goes on to claim that drives and affects work together:

> by associating with an affect, a drive acquires sensitivity to a stimulus and thereby 'knows' when to activate; conversely, an affect can give better shape

> to its pattern of behavioural response by taking up a pursuit object from a drive.
>
> (2012:14)

Here we get a sense of the mutual recruitability of drives and affects. It is this idea of recruiting which Anderson suggests explains ambiguous attitudes. He uses the example of an artistic drive to destroy: should this particular drive recruit or be recruited by the affect of a 'Dionysian overflowing of joy and superabundance' (2012:15) then this is a fruitful relationship conducive to flourishing. Conversely, if the same drive recruits or is recruited by a vengeful affect such that it expresses a hunger, then this is a relationship which is inimical to flourishing. Actions are deemed conducive or inimical to flourishing by virtue of the configuration of the drive/affect interaction.

Anderson then sketches the model of drives and affects in more detail, proposing that the 'affects form a cross-hatched, mutually supporting structure of attitudes, whose integration rests on the way they are structurally tailored to recruit one another' (2012:15). He also specifies that there exists a 'one-many' relation between drives and affects, whereby the same affect can be recruited by many different drives, just as the same drive can be recruited by various affects. This image of a 'cross-hatched' structure with one-many relations between drives and affects via recruitability, highlights a crucial difference between Nietzsche's account of the self and Hume's bundle theory of the self, since it does not appear that drives are 'completely loose, distinct existences' (2012:16). There is some shape to this structure. Anderson also suggests that this structure is not a 'mere 'stage' upon which they enter and exit for one-off causal interactions. Instead the self should be seen as a '*repository* of recruitable drives or affects that are always **(p.224)** available to complete any of its given active drives or affects' (2012:16). Moreover, several drives from the repository may combine to create more complex drives.

It is this notion of self qua repository which leads Anderson to claim that Nietzsche allows a 'minimal self' to emerge. Anderson argues that although the minimal Nietzschean self is 'built out of the drives, affects, and other attitudes, and could not be what it is without them', similarly 'the drives and affects could not be what they are without the whole

Nietzschean self either, in that, for example, the typical complements and contents, and hence the functional capacities, of a given attitude will depend on which *other* drives and affects are available for it to recruit' (2012:16) – hence the idea of the self as a repository is required.

He suggests that this is further supported by the fact that many of these drives and affects, far from being 'fleeting occurrent states *à la* Hume', actually persist over time (although as previously mentioned, this is not to say that drives and affects cannot be acquired or lost, since the boundaries of the minimal self are porous). Since the drives/affects generally persist, Anderson argues that the minimal self must have its own:

> *separate*, diachronic identity, which persists across *changes* of drives and affects. The minimal self is thus but one psychological structure among the others. It acquires the right to the name 'self' simply in virtue of being the emergent structure that encompasses *all* of the substructures **(p.227)** available for recruitment by one another.
>
> (2012:18)

This model of the minimal self provides Anderson with an interesting way of accounting for responsibility in the Nietzschean self. He achieves this by claiming that 'the self—qua the emergent structure encompassing *all* the co-recruitable attitudes—can suffer from a 'gap' between its own activity and that of some constituent(s)' (2012:18). Although drives and affects are recruitable, this does not guarantee that they will be successfully recruited in the appropriate circumstances.

In such a case, the drive in question remains part of the totality, in that it still exists within the repository that is the minimal self. In this way, Anderson claims:

> the minimal self can remain 'responsible' for a recalcitrant attitude as something that belongs to it—by contrast to a mastering drive or a 'crystallization', wherein any attitude that is not presently and actually functionally integrated is simply not a part of the emergent whole, but a separate factor.
>
> (2012:18)

Here, Anderson is able to preserve some element of the personal, and of ownership. The question remains whether this is a satisfying conception of responsibility. It may be helpful at this point to recall Gemes' distinction between deserts free will and agency free will: if we assume that Nietzsche accepts agency free will, rather than deserts free will. then perhaps this notion of responsibility which Anderson advances in his account seems appropriate.

With this idea of the minimal self, we are left with a conception of responsibility where notions of punishment and reward (required for deserts free will) are not applicable, or appropriate. Instead all we have is a more detailed description of what occurs during the self during agency: some drives or affects may or may not be recruited successfully in the appropriate circumstances, and this will be dependent upon the various other configurations of drives/affects operating at the time and the external circumstances. Any sense of free will is purely in the minimal self's capacity qua emergent psychological structure to stand apart from its configuration of drives and affects.

It is this which allows the self to:

> have the capacity to take up attitudes (including evaluative attitudes), towards the world, and also towards itself and its drives and affects. These reflexive attitudes may include consciously reflective, or even deliberative attitudes, such as the control of affective interpretations involved in perspectivist objectivity, or the more or less explicit attitudes of self-management involved in Nietzschean self-overcoming, self-mastery, and so on.
>
> (2012:19)

So how should we make sense of self-formation?

Anderson is however careful to note that:

> the minimal self *is just* a certain moral psychological structure among the drives and affects, no matter how conflictual and disunified they may be' and maintains that 'one must attain something further to become a self in the stronger, normative sense.
>
> (2012:21)

Anderson here differentiates between two manifestations of the self: 1) the basic minimal self and, 2) the normative self which has been created and realises a form of autonomy. Can we progress from one to the other, and if so how can this be done?

Anderson notes that the minimal self is not unified:

> The degree of unity possessed by the minimal self is limited, not only in that drives and affects may be unavailable to central consciousness and completely non-transparent to one another, but also in that **(p.225)** different constituents of the self may stand in oppositional, even quite conflictual, relations, resulting in weakness of will, and the like.
>
> (2012:17)

The normative self by contrast cultivates some unity, specifically self-generated unity:

> What makes such unity count as *one's own* is precisely its having been *self-generated*—that is, the unity among my drives and affects arises from regulating control over them that is exercised by and through the attitudes proper to the emerging self: to be noble is 'to have and not have one's affects…at will; to condescend to them…to make use of [them]…' (*BGE* 284).
>
> (2012:22)

Here, Anderson claims that this is merely an ideal for the relation that ought to obtain among the drives and affects, and that 'whether any individual person attains that ideal or not is an empirical question, to be settled by the best interpretation of the person's life' (2012:22). Self-formation on this account would seem to rely somewhat on external factors, which would influence the repository of drives and affects, which then go on to mutually recruit each other in ways which are conducive to flourishing, and which would generate a more 'unified' structure.

But what should we make of Anderson's conception of unity? Here we see that Anderson assumes that transparency of drives and affects is vital for unity, along with the idea that they must be available to a 'central consciousness' (he does not, however, go into detail about how exactly this central consciousness is to be understood and what role it has in

the self[9]). He also makes the assumption that unity is dependent upon synthesis, and suggests that the relations between drives and affects cannot be one of synthesis, since this would imply that said synthesis is achieved by some separate, unified agency (drawing us dangerously close to the Kantian picture). However, as Janaway suggests in his account we would do better to think of *Einheit* in Heraclitean terms where no notion of synthesis is required yet there is still unity.

Anderson thus characterises the self as plastic and capable of creative agency. With his notion of recruiting, he illuminates the interactions between the drives and affects, and brings in the notion of diachronic identity. The distinction between a minimal self and a normative self means that self-formation is something to be achieved, but Anderson does not limit this possibility only to the 'Sovereign Individual'.

1.4 Evaluation of the Debate

Having examined some of the interpretations of the self in Nietzsche it seems that in no way are the various accounts equivalent,[10] and this largely due to the fact that they are motivated by different concerns. These concerns range from:(i) normativity, (ii) agency, (iii)autonomy and freedom, (iv)cognition, (v)conscious willing, (vi)self-knowledge, (vii) self-governance, and (viii) first-person subjectivity.

The extent to which commentators pay attention to the specific workings of the components of a Nietzschean self seems to vary. Accounts such as Gemes' tend to focus more on the possibility of autonomy and freedom than the inner workings of the self:

[9] All that is said regarding consciousness in Anderson's account pertains to his assertion that the self for Nietzsche cannot be transcendental: 'moreover, the boundaries of the minimal self, unlike those claimed for the transcendental ego, are not identical with those of consciousness. In fact, the boundary mismatch obtains in both directions: the minimal self encompasses drives and affects it is not aware of, and it may have apparent conscious awareness of powers (e.g. the will) that are illusory. Thus, there can be no a priori argument from the alleged unity of consciousness to a strong, transcendental unity proper to the minimal self.' However, Anderson does not clarify the status of consciousness in his account of the minimal self; it is not clear whether consciousness exists in the repository at the same level of the drives and affects or whether it exercises any influence on the process of recruitment.

[10] In the following groupings and comparisons (see table) I wish to only point out general trends amongst the commentators not to say that in each account the commentator does not discuss x, y, or z at all or that he only discusses x, y, z.

whilst these topics certainly overlap, it is not clear that Gemes pays sufficient attention to the actual mechanics and make-up of the self. Similarly, neither Gardner nor Leiter offer any detailed treatment of the psychological apparatus Nietzsche provides. Katsafanas, Janaway, and Anderson by contrast, look at the operations of within the self in great detail, and attempt to trace the exact workings of these psychological elements. But only Janaway and Katsafanas attend to the cognitive aspects of the self.

Commentators also seem to work with different understandings of the following notions: (i) free will (incompatibilist notions versus compatibilist notions), (ii) consciousness (consciousness qua awakeness/ awareness/ substantive faculty/ reflection/ or self-consciousness), and (iii) the self (qua state versus qua achievement). The definitions adopted can greatly influence accounts: for example, taking consciousness to be bare awareness accords consciousness a far more passive role than if we viewed it as a reflective capacity. Not all of the accounts discussed make it clear what they take these notions to be in Nietzsche's works, or they offer very selective readings which ignore other alternative interpretations, e.g. Leiter on consciousness or free will.

A running theme in many of the accounts (Gemes, Janaway, Katsafanas) is the idea that there are two levels of the self: 1) the self qua state and, 2) the self qua achievement i.e. a 'better' self. This distinction introduces a normative dimension and directly speaks to the notion of self-formation. It is not always clear, however, how we ought to go about progressing from one to the other: neither Gemes nor Anderson, for example, provide much information as to how this may be achieved.

Comparing and contrasting the accounts examined so far in the table below is helpful to see the amount of variation and weighting of concerns in these interpretations:

	Static or Plastic	Detailed discussion of drives and affects	Consciousness as epiphenomenal	Naturalistic or Transcendentalist or other	Self as task or achievement	Freedom, Autonomy possible
Leiter	static	no	yes	naturalistic	no	no
Gardner	n/a	no	no	transcendentalist	no	yes

Katsafanas	plastic	yes	no	other	no	n/a
Janaway	plastic	yes	no	other	yes	yes
Gemes	static	no	yes	naturalistic	yes	yes
Anderson	plastic	yes	n/a	other	yes	yes

Conclusion

We can see that various models of the self emerge, which can be radically different, and alongside them, varying understandings of what self-formation might consist in. At times, the debate in the secondary literature appears somewhat unclear– perhaps this confusion would be avoided if we had a clearer sense of what self-formation entails. Each model seems to assume a different understanding of self-formation. Sometimes it looks like the process should culminate in an individual who is: (i) free and/or autonomous agent, (ii) displays a highly aestheticized notion of *Einheit*, (iii) capable of self-governance, (iv) possesses self-knowledge, (v) possesses perspectival knowledge, or (vi) capable of self-affirmation. It is not obvious what Nietzsche had in mind with the idea of self-formation; whether this process privileges one of the above characteristics, or involves a combination of these features.

Reconstructing a model of a Nietzschean self would be impossible without detailed discussion of the drives and affects. However, it seems that we also require a sustained discussion of the other aspects of the self, which Nietzsche discusses, e.g. memory, imagination, forgetting, intellect, to provide an exhaustive analysis of a Nietzschean self. It is not obvious that any of these components can be reduced to a drive or an affect. To leave these features out of an account of the Nietzschean self would be somewhat remiss.

The rest of the thesis will explore some of these components, and note that, from Nietzsche's remarks, they appear to have distinctive regulatory functions. As such, the idea of self-regulation may be a useful notion to employ in exploring Nietzsche's treatment of the self. The idea of self-regulation may be helpful in answering the wider question: what kind of self can engage in self-formation? One answer is that since Nietzsche presents self-formation as a particularly challenging task, the kind of self that

would be capable of such a feat is one which is relatively stable and resilient. The thesis will explore the mechanisms deployed in a Nietzschean self to generate these necessary conditions for self-formation.

2. Nietzschean Forgetting and Reinterpretation

Introduction

This chapter aims to elucidate Nietzsche's understanding of forgetting,[11]and in doing so, sheds light on different structures within the Nietzschean self.[12] I suggest that forgetting qua capacity has a different functional profile than that of a drive, and that this capacity emerges as a crucial regulatory process within the self, providing a necessary condition for value generation and reinterpretation.

Nietzsche presents what seems like an idiosyncratic understanding of forgetting. Contrary to the notion of forgetting as a passive, degenerative process i.e. poor memory-performance, associated with decline and incapacity, Nietzsche offers us a far more nuanced, and (more importantly) positive, understanding of forgetting: 'it is altogether impossible to *live* at all without forgetting' (UM II:62).[13] The comments he makes regarding forgetting are prima facie paradoxical as with the notion of active forgetting in the *Genealogy*: 'forgetfulness is not just a *vis inertiae*, as superficial people believe, but is rather an active ability to suppress, positive in the strongest sense of the word' (GM II:1), or puzzling as in 'On Truth and Lies in a Non-Moral Sense', where Nietzsche claims that we form concepts like 'leaf' by '*forgetting* the distinguishing factors' [*durch ein Vergessen des Unterscheidenden gebildet*]. To what degree can we understand this sort of filtering as a case of forgetting?[14] It would seem to run counter to our commonsense understanding of forgetting: one would presume that we can only forget that of which we have had first-person awareness: in other words, we might tend to limit forgetting to cases where someone, was for a time at least, in a position to remember. However strange or unclear

[11] I do not have the space here to discuss the broader issue of exactly how Nietzsche's views on forgetting develop across his works. For the purposes of this chapter I shall focus on his remarks on forgetting which seem particularly pertinent to the issue of self-regulation, drawing from both early and later works.

[12] Forgetting has hitherto been neglected in accounts of a Nietzschean self by commentators. Even Stern, despite offering perhaps the most exhaustive taxonomy of a Nietzschean psychology in his 2014 paper, noticeably does not discuss forgetting (not even under 'miscellaneous' (2014:9)).

[13] This is not to say that I wish to attribute to Nietzsche a denial of forgetting as understood in common parlance. This chapter aims to show that Nietzsche offers an additional rather than alternative understanding of this phenomenon.

[14] Here we see Nietzsche's broader understanding of forgetting. It is beyond the scope of this chapter to discuss this in full, but in brief, the use of forgetting as filtering in 'On Truth and Lies in a non-moral sense' relates to what other philosophers (e.g. the British Empiricists), term 'abstraction.' Nietzsche suggests here that 'every concept comes into being by making equivalent that which is non-equivalent' through 'forgetting' particularising features, and that this gives us the illusion that reality is more stable than it actually is. He suggests that this filtering imbues our cognitive processes in general, and strategically creates psychologically valuable perspectives for the individual.

these remarks may be, it seems important to examine forgetting as Nietzsche frequently singles out this phenomenon in his works and stresses its importance: 'a little *tabula rasa* of the consciousness… active forgetfulness, like a doorkeeper or guardian of mental order, rest and etiquette from which we can see there could be no happiness, no hope, no pride, *no present*, without forgetfulness' (GM II:1). How should we may sense of this array of remarks?

An initial objection may be that we ought to be cautious about conferring too much importance upon this notion of forgetting, and its implications for the self: in the *Genealogy,* Nietzsche presents this phenomenon as part of a larger story about the emergence of memory as a capacity required for making and keeping promises (which is, in turn, a presupposition of responsibility). This memory-based ability to promise is then aligned with, for example the Sovereign Individual and an ideal of free agency (Gemes; Poellner; Richardson 2009). However, the status of the Sovereign Individual is somewhat fraught (cf. Acampora 2006[15]), which may lead us to doubt the importance of a memory-based capacity to make promises. I will remain neutral on this particular debate for the purposes of this chapter but will merely note that even before considering that question it is necessary to ascertain Nietzsche's stance on forgetting. Moreover, given his comments on forgetting throughout his work, it seems that this phenomenon merits attention in its own right.

Examining Nietzschean forgetting will bring up such questions as: how can we understand this characterisation of forgetting as active? And just how different is this to Freudian repression? I suggest that Nietzschean forgetting is construed as 'active' in the sense of being an operational subconscious process. Rather than being viewed as a derivative or secondary effect that occurs due to the malfunctioning of some other mechanism such as a memory faculty, it is a capacity in its own right. Crucially, it differs

[15] Acampora (2006) suggests that viewing the Sovereign Individual as one of Nietzsche's ideals is potentially problematic. Aside from there not being very much textual evidence in support of the Sovereign Individual (152-153), the very notion does not seem to cohere with what Nietzsche writes about subjectivity (specifically the composite nature of the self) in passages like BGE 16-20 and GM I:13. Moreover the emphasis on autonomy and free will which we find in the Sovereign Individual potentially rings Kantian alarm bells.

from the Freudian account in that it involves (at least in its ideal state) a total metabolisation of affective material rather than a temporary suppression (where the material lies latently and can re-emerge via sublimation). It is this metabolisation which accounts for the plasticity of the self. This sense of forgetting, I suggest, is an idealised state in the self. Not everyone will be capable of achieving it - Nietzsche seems to restrict such a forgetting to the 'strong' or 'healthy' types. Even those individuals who are capable of this forgetting do not consistently achieve this idealised state.[16]

Here is the plan for what follows. Section One addresses an objection to the Nietzschean concept of 'active forgetting', namely that it is prima facie paradoxical. I propose that by understanding forgetting as an operative subconscious capacity akin to a faculty,[17] (we may compare this to a language faculty, functioning but beneath our conscious control), Nietzsche provides us with a philosophical psychology which does not in fact come up against this paradox. Section Two examines the functional profile of forgetting, arguing that forgetting does not fit the criteria of a drive outlined by commentators such as Anderson (2012), Richardson (2004), Katsafanas (2005, 2012, 2016), and Janaway (2009), and that it is instead better understood as a capacity. I suggest that this capacity is responsive to, and recruited by, the drives in order to reach their ends.[18] Section Three examines the most important role forgetting has within the self as a regulatory 'deep-clean' procedure, and examines precisely why Nietzsche suggests this mechanism is required. This section also discusses how the Nietzschean account of forgetting differs from the Freudian model. Forgetting frees up space in the mental economy via alterations of the affective valence[19] of memories. Such space is vital for the re-interpretative process

---

[16] There is a clear albeit difficult to delineate distinction between individual and collective forgetting (as seen in for example *On the Genealogy of Morality*, where Nietzsche discusses grand cultural forgetting and presents the genealogical method as an act of remembering and a means of gaining self-knowledge). For the purposes of this chapter I will focus only on individual forgetting.

[17] Bearing in mind Nietzsche's suspicion of 'faculties' (BGE 11) I do not suggest that we straightforwardly call forgetting a faculty, but it is a useful way of distinguishing this mental process from something like a drive.

[18] The idea of drives recruiting has been discussed by Anderson (2012) and noticeably does not entail that we regard the drives as homunculi. Katsafanas (2005) also convincingly argues against the idea of drives as homunculi.

[19] There is of course a wide debate in the philosophy of emotion as to how best to understand the notion of 'affective valence.' For the purposes of my chapter I will take 'affective valence' to mean the positive or

and allows new values to displace old ones, something which would underpin a phenomenon like self-formation. This section also explores the importance of this revised model of Nietzschean forgetting for the understanding of several other Nietzschean themes such as the Eternal Recurrence and Amor Fati, both key components of his practical philosophy.

2.1 The 'Paradox' of Active Forgetting

Prima facie, the notion of 'active forgetting' seems paradoxical. If I am 'actively' forgetting, surely I would be undermining my attempts to forget in continually trying to achieve this. We might compare this paradox to conscious self-deception or planned spontaneity. Such activities defeat their very purpose. Some commentators characterise active forgetting in such a way that this paradox does rear its head. Blustein (2014:106) states that for Nietzsche 'who sees volition in both remembering and forgetting…active forgetfulness is a reflective and purposive effort'. The term 'reflective' immediately has connotations of some central controller orchestrating the process. The status of reflective thought for Nietzsche is in any case, contentious; there are passages where he seems to affirm and even exalt reflective thought (D 41), but equally passages where he doubts its very possibility (BGE 3; D109).[20] But is 'reflective' the only way to understand 'active'?

Let us examine the passage in detail:

> Forgetting is not merely a *vis interiae* as the superficial imagine; it is rather an active and in the strictest sense, positive faculty of repression, that is responsible for the fact that what we experience and absorb enters our consciousness as little while we are digesting it (one might call the process

negative character of memories or an aspect(s) of memories. The amendment of valence in these cases would most likely be a quantitative one, i.e. a lesser degree (amount) of emotion x.

[20] See Katsafanas (2012,2016) and Leiter (2002) for opposing discussions.

> "inpsychation" [*Einverseelung; Einverleibung*[21]] as does the thousandfold process, involved in physical nourishment- so called "incorporation".
>
> (GM II:1)

It is not obvious that we need to characterise 'active' as 'reflective' here. In fact, the passage above suggests that this active forgetting is something separate from consciousness (where we would normally locate reflective thought), since it controls and is 'responsible' for what enters, remains, and exits our consciousness.

So how else may we understand 'active' forgetting?

The word Nietzsche uses is *Hemmungsvermögen.* A *Vermögen* may refer to a power, an ability, a disposition, or a capacity. The literal translation of the passage here is that this inhibition capacity actively inhibits the digestion of our experiences from rising to the surface of consciousness. Whilst we may be accustomed to thinking that we can consciously manipulate and direct such things, Nietzsche makes clear that many of our mental processes operate covertly at a subconscious level, and that we tend not to have any control over them. In this way, an active mental process motivates our actions but in such a way that it is not within our conscious control: we can compare this to subsystems in the body - circulatory, respiratory etc. which operate automatically. Active is thus understood as automaticity in that it is: a) not consciously willed, and b) not something that the conscious self can prevent.[22] In the next section, I go on to suggest how this process is guided by the drives such that the forgetting capacity is active when it is enlisted or 'in use' by the drives. The psychology we are provided with allows us to call forgetting active and enables us to bypass the paradox outlined previously; forgetting is understood as the activity of a mental process which operates at a subconscious level where we cannot choose to initiate the process or decide its target.

---

[21] *Einverseelung* meaning 'making it part of the soul'; *Einverleibung* meaning 'making it part of the body.'
[22] This is not to say that active always equals automaticity or that we cannot have active processes which are not automatic.

In GM II:1 Nietzsche offers a characterisation of forgetting as a capacity in its own right. Rather than viewing forgetting as a malfunctioning of memory, i.e. as secondary or derivative, Nietzsche refers to forgetting as 'an apparatus of suppression' which can be 'damaged, so that it stops working'. Here, it seems that forgetting is a stand-alone mechanism which can potentially malfunction or be impaired in some way – the person whose forgetting mechanism is damaged is compared to a 'dyspectic' who 'cannot 'cope' with anything'.[23]

Forgetting as an active capacity may thus be understood as purposive, but not conscious or reflective. In this way Nietzsche bypasses the paradox (and thereby does not need to offer any solution to it). The paradox is only relevant if forgetting is characterised as something we can consciously choose to put into practice and direct.

In the next section, I will explore in greater detail why forgetting for Nietzsche does not fit the criteria of a drive, and how this capacity interacts with the drives.

2.2 How to Understand Nietzschean Forgetting

[23] Striking comparisons may be made here with contemporary neuroscience, and specifically the theory of 'motivated forgetting.' Michael Anderson (2012, 2014) details how mechanisms of inhibitory control exist within the mental economy, specifically the suppression of intrusive thoughts via hippocampal activity. Such activity prevents memory retrieval and is common in trauma victims who experience memory deficits. In this way, forgetting, though not experienced at a conscious level, is far more *active* than we realise; i.e. it is an operational process within the mental economy. With regards to the idea that, as an independent capacity, this function can be damaged or impaired in individuals, studies have noted that those suffering from post- traumatic stress disorder, exhibit insufficient memory suppression. Here it looks as though this type of forgetting has a specific purpose, i.e. to limit trauma.

The previous section suggests that forgetting is best understood as a capacity.[24] Drive-theory accounts, however, may try to account for the active nature of forgetting by characterising it as a drive: Hales and Welshon (2000:168), for example, refer to 'the drive to forget'. However, I argue that forgetting does not fit the criteria of a drive outlined by commentators such as Katsafanas (2005, 2009, 2016), Anderson (2012), Janaway (2009), all of whom characterise drives as being: a) evaluative, b) directional, and c) responsive to other drives and affects. Prima facie it may be tempting to include forgetting in this but there are important differences.

Firstly, drives tend to operate in a particularly specific and targeted manner. Katsafanas (2005, 2016) refers to Freud's distinction between the aim [*Ziel*] of a drive and the goal or target [*Objekt*] of a drive. Take, for example, a drive to aggression – the aim of this drive is just to express aggression but in order to manifest this activity, the drive needs to have an object: e.g. in heavy traffic, the object of someone's drive to aggression might be a slow cyclist. We can make sense of the idea of a blanket aggression finding specific targets. It seems difficult to map this distinction between an aim and a goal or target on to forgetting. If the aim is 'to forget', then what is the scope of this aim? It seems somewhat extreme to posit some mechanism whose aim is to forget everything. And if we want to weaken that aim, and say the aim of the drive to forget is to forget certain things but not others, it is not clear what would underpin such limits or rules regarding what is to be forgotten. In order to achieve anything as sophisticated as selective forgetting, content would have to be evaluated in some way to ascertain whether it should be retained or actively suppressed. To say that a drive to forget is sophisticated enough to choose precisely what to forget brings us to the paradox of forgetting. It therefore seems problematic to say forgetting has an aim in the sense where calling it a drive proves to be explanatory. The functioning of forgetting would therefore be better aligned with a

[24] It seems odd to characterise forgetting as any of the following components available to us in a Nietzschean model: an emotion, an affect, a *Hang* - in the sense of a 'taste' for x – (BGE 2, GS 87) or an instinct.

capacity which does not operate in as sophisticated a manner as a drive, i.e. it does not need to select and aim at specific targets.

Nietzsche refers to conflict between drives, where each drive seeks to dominate and come out on top. A drive must therefore at least potentially be able to occupy the seat of control amongst other drives. What would a self look like where the drive to forget dominates? This seems an unlikely picture: we are left with the amnesiac (in which case this would be explained as a secondary effect of the memory faculty being damaged). Even if we posit a scenario where it looks as if a drive to forget is dominating (e.g. with someone drinking excessively to forget some trauma), then this is explained via other drives or affective states (e.g. a drive to shame, disgust etc.). If we then want to suggest that the drive to forget is just a drive which necessarily plays a subordinate role to other drives, this is also problematic. No other drive seems to be such that it only ever exists in a subordinate relational structure, making it difficult to defend the claim that the drive to forget has its own distinct structure.

The last point of difference between a drive and forgetting is the way in which the activity of both can be impeded. For the activity of a drive to be inhibited, it must be repressed so that other stronger drives dominate and the operation of weaker drives is restricted. In D109 Nietzsche describes how we may combat or train drives so as to cause them to 'wither away'. However, in such cases it seems all that happens is that the drive exists latently. To continue the horticultural metaphor Nietzsche uses in D560 we might compare this to a latent flower bulb. The bulb itself is not impaired or damaged: it is rather that conditions for its flowering are not present. Similarly, latent drives are not impaired or damaged – they retain (if only potentially) their functionality. A drive is therefore not damaged; rather it is restricted by a stronger drive. Forgetting however seems to be a difference case. As discussed in the previous section, Nietzsche makes it clear that our ability to forget can be damaged: 'this apparatus of suppression [*Hemmungsapparat*] is damaged so that it stops working' (GM II:1). In this way forgetting seems more akin to a faculty in that it can be impaired: we might compare this to the case of language (usually characterised as an ability or faculty), and something like Broca's aphasia or Wernicke's aphasia where specific areas of the brain are damaged resulting in

a loss of this capacity. It is not that our ability to construct and operate language is repressed, but rather that this ability is damaged.

For these reasons, it seems we ought to resist characterising forgetting as a drive. Conceiving of it as a capacity allows us to: 1) avoid the paradox of active forgetting since it operates outside of conscious thought, and 2) account for a weak sense of an aim without having to regard forgetting as evaluative and directional.

This brings us to the question of the relationship between this capacity and other components of the mental economy, and specifically the drives and affects. I suggest that in examining this relationship, we can come to understand the process of forgetting in the self, i.e. how it is that certain material is forgotten and why.

To explain this, I will draw upon Anderson's idea that psychological components within the Nietzschean self possess the ability to *recruit* one another. Anderson (2012:14) suggests that drives and affects are mutually recruitable: 'by associating with an affect, a drive acquires sensitivity to a stimulus and thereby 'knows' when to activate; conversely, an affect can give better shape to its pattern of behavioural response by taking up a pursuit object from a drive'. For example, I might be very hungry as I stand waiting in a long queue to pay for my sandwich; if someone tries to jump the queue, my drive for food might recruit my affect of anger and I snap at the offender, thus ensuring my drive for food reaches its object. With the following scenarios, I suggest that the forgetting capacity can also function in this manner, i.e. it is sensitive to affective states and responsive to drives, and recruited by them to achieve their ends/to reinforce and sustain the interpretations they generate.

In GS 127 Nietzsche gives the example of different perspectives or interpretations of a 'violent stimulus', claiming that whether it is 'experienced as pleasure and pain is a matter of the interpreting intellect, which, to be sure, generally works without our being conscious of it; and one and the same stimulus can be interpreted as pleasure or pain'. We can imagine that in such a case, these interpretations would need to be reinforced in some way such that they maintain sufficient weight for the individual. To give a specific example: running a marathon is at times an undoubtedly painful experience, yet many long-distance runners will report deriving joy and satisfaction from such an event, and

repeatedly sign up to races. Katsafanas (2012) also uses the example of pain during exercise, suggesting that it is rendered bearable or even pleasurable by coupling the affective experience with the consciously generated interpretation 'this pain means that I am exercising vigorously'. This certainly seems plausible: particularly within the context of a race, runners might attend to one or maybe more positive interpretations to motivate themselves: 'this pain means I am exceeding my limits', 'this pain indicates I'll achieve a new personal best time', 'this pain means I must be a frontrunner in the race' and so on.

The idea of interpreting sensations like pain is particularly interesting when we consider retrospective interpretations, i.e. how a marathon runner evaluates a race after the event. This gives us a good example of how the forgetting capacity can be used to reinforce positive interpretations (generated or sustained by e.g. a drive to win, and affects like pride), and to allow the individual to avoid placing too much weight on negative affective valence, i.e. remembering the degree and intensity of pain experienced during the race, and any associated negative interpretations. This kind of explanation may go some way in explaining why it is that runners continue to participate in marathons.

Babel's (2015) study[25] demonstrates that runners do not accurately remember the pain induced by physical exercise. When these marathon runners recall a race, we might well speculate that something like the capacity of forgetting would be in operation: the marathon runner would be able, through this capacity, to dial down any negative affective valence i.e. remembering the degree and intensity of pain at the time. If he were able to recall this accurately, this could potentially cause him to avoid signing up to races in the future. Moreover, remembering this negative affective valence might give negative interpretations he held at the time more prominence e.g. 'I felt like I wasn't cut out for marathons', and thus discourage him from running again.

[25] This study consisted of 62 participants who ran the eleventh Cracovia marathon in 2012. Babel gave these participants questionnaires a) straight after crossing the finishing line, and b) three or six months later to gauge the following: (i) intensity of pain experienced, (ii) unpleasantness of pain, and (iii) their feelings of positive or negative emotions. Using a 7 point scale to measure the level of pain, on average, the participants reported a level of pain of 5.5 after they had just finished the race. This then dropped to an average of 3.2 regardless of how much time had passed (three months as opposed to six months).

In this way, forgetting would be responsive to other drives, such as the drive to win, and other affects such as pride (at winning, in one's sporting ability), and shame (at losing and not meeting one's standards).[26] In such a scenario, both positive and negative interpretations related to the event may be retained, but the negative affective valence is dialled down.

We could also see drives recruiting the forgetting capacity in scenarios where the material which is deemed inimical to the goal of a drive is *external* to the individual (unlike the case of the marathon runner where the material is internal to his own mental apparatus). Katsafanas (2012) refers to the drive to jealousy. This is an interesting example as it inevitably involves both an interpersonal element and external factors. Under the influence of this drive an individual is inclined to give excessive weight to certain features of their environment[27] and interactions, e.g. glances, tones of voice and so on. However, it is not just that the individual, as Katsafanas suggests, sees justifications 'for jealousy everywhere' or is 'drawn to certain features of his environment that another agent would scarcely notice' (2012:12).

The picture is more complicated. A jealous individual not only attributes greater weight to features of his environment but simultaneously filters out or suppresses or *forgets* any reason which would counter this jealousy or other elements in his environment which would not support his suspicions. For example, a jealous husband tracking his wife's movements may panic at her taking a route which involves passing an ex-boyfriend's house, and forget the fact that this route is the quickest, most convenient route to her new workplace (a fact he knows very well). Even if this piece of information is recalled by the husband it will be filtered out as not being a viable explanation of his wife's behaviour. In this way, the capacity to forget is recruited to subdue certain factors and enable others to

[26] Understood in this way, we may, as Blustein (2015:108) notes, draw a parallel between forgetting and Hursthouse's notion of 'multi-track' dispositions which are: concerned 'with many other actions as well, with emotions and emotional reactions, choices, values, desires, perceptions, attitudes, interests, expectations and sensibilities.'

[27] We might compare this to the notions of 'mood' and 'emotion congruence' in the philosophy of emotion, which suggest that affective states tend to draw attention to specific aspects of the surrounding environment, sometimes resulting in detrimental epistemological effects for the individual. My thanks to Fabrice Teroni for pointing out this link.

come to the fore via the operations of the drive to jealousy. The capacity of forgetting is thus able to interact with and be recruited by other drives, forming a subordinate relational structure.

In the previous section, I suggested that in GM II:1 Nietzsche presents forgetting as a power [*Kraft/vis/Vermögen*] that inhibits the becoming conscious of the processing of an experience. In this way, Nietzsche's understanding of forgetting is to be distinguished from a momentary incapacity to recall x. In this section, I have attempted to offer a more detailed explanation of this psychological mechanism. In the examples above, we have instances of this capacity of forgetting as selective experiences.[28] Here we can see how the capacity of forgetting interacts with drives and affects to act upon specific material, and that this material can be either internal or external to the individual.

We now have some idea of what kind of psychological structure forgetting is, and how it functions and interacts with other components of the mental economy. The purpose of this chapter is to elucidate how forgetting acts as a form of self-regulation. The next section will examine in closer detail the idea that forgetting as unconscious automatic processing, (if it functions correctly), changes affective valences of remembered experiences.

2.3 The Role of Forgetting in Self-Regulation

This now brings us to the role forgetting plays in self-regulation. I will explore: 1) why Nietzsche thinks forgetting is required for this purpose, and 2) exactly how forgetting contributes to self-regulation. I suggest that in amending the affective valence of memories, forgetting: a) preserves the strength and health of the mental economy, and

[28] We might compare this to the notion of a 'confirmation bias' in psychology. This is also similar to Riccardi's point in 'Nietzsche's Sensualism' (2011:246ff) about selective perceiving. Using Papineau (2007), Riccardi argues that selective perception can be seen as using sensory templates. There is a parallel here with the idea of forgetting as filtering where things are interpreted in light of a hypothesis – in the example of the jealous husband the filter is 'my wife is cheating'.

b) creates new mental space needed for value generation. In this way, forgetting provides a necessary condition for Nietzsche's practical philosophy, specifically the idea of reinterpretation.

Contrary to the Platonic emphasis on reminiscence and recollection, Nietzsche proposes that forgetting is an important component of healthy cognition and a *conditio sine qua non* of flourishing, repeatedly aligning forgetting with strength: 'his strength lies in forgetting himself' (UMII: 4; cf. GM I:10, GM II:1, EH 'Wise' 2). Nietzsche is almost unreservedly positive about the phenomenon of forgetting,[29] and this psychological mechanism, according to commentators such as Swanton (2015) and Blustein (2014), is even elevated to the status of a virtue.[30]

For Nietzsche, a key characteristic of the strong and healthy is that they are able to 'forget' in the sense of 'moving on from x' or 'not dwelling on x.' 'X' here stands for experiences which one has found to be unpleasant (including one's own misdemeanours). Why does Nietzsche advocate this process of moving on from x or not dwelling on x? In order to understand the full weight of this, it is worth briefly examining Nietzsche's idea that we need to have the right kind of relationship to the past.

What would this relationship look like? We can begin to sketch out an answer to this question by looking at examples where Nietzsche seems to detail problematic relationships with one's personal past. In works such as the second *Untimely Meditation* and *On the Genealogy of Morality*[31], there are passages where Nietzsche displays a clearly

[29] In the Preface to *The Gay Science* Nietzsche even aligns forgetting with the artistic: 'oh how we now learn to forget well, and to be good at *not* knowing, as artists!' and contrasts it with the less admirable 'will to truth' which seeks to uncover and unveil what should remain hidden.

[30] As Swanton suggests, forgetting is viewed as being 'a virtue of considerable psychological importance which also underlines other virtues like justice' (Swanton 2015: 161), and in this way, it appears as a virtue which possesses both intrinsic and foundational importance. Blustein (2014:107) also refers to forgetting as a Nietzschean virtue.

[31] This is not to say that Nietzsche is wholly negative about memory in these works. In the second *Untimely Meditation* and his discussion of critical history, Nietzsche appears very positive about the use of memory vis-à-vis culture and history at some larger level. In the *Genealogy*, the very method of genealogy appears as a prodigious exercise in memory, which Nietzsche clearly advocates.

ambivalent or even negative attitude towards the idea that our past plays a role in shaping us.

I suggest that two characteristics to be avoided emerge: 1) an over-preoccupation with these memories (regardless of whether the memories are good or bad) as seen at the beginning of the second *Untimely Meditation*, 2) a tendency to utilise negative affective content from unpleasant memories to facilitate phenomena like *ressentiment*, as seen in the *Genealogy* with the example of the slaves.

If an individual displays this kind of relationship with their past, the kind of self which is then constituted allows past events too much influence on their emotional lives, their attitudes, and their actions. Given the emphasis Nietzsche places upon reinterpretation, it seems that this kind of relationship with one's personal past is problematic.

Let us look at this idea of over preoccupation with the past in closer detail. Whilst much of the second *Untimely Meditation* is concerned with the past qua history or culture on a grand-scale, certainly at the beginning of this work, I take many of Nietzsche's remarks to apply to the past of an individual. Here, Nietzsche first presents the idea that we can, at times, be overly-preoccupied or overly-dependent upon the past, and that this can occur regardless of whether the content of our memories is positive or negative. A consequence of this over-preoccupation with the past is that it may impede present and future projects, thus debilitating the individual. In this way, autobiographical memory has the potential to be a psychological burden – Nietzsche describes it as a 'chain':

> He also wonders at himself, that he cannot learn to forget but clings relentlessly to the past: however far and fast he may run, this chain runs with him. And it is a matter for wonder: a moment, now here and then gone, nothing before it came, again nothing after it has gone, nonetheless returns as a ghost and disturbs the peace of a later moment.
>
> (UM II:61)

Here Nietzsche focuses on the relationship one has to the past and one's memories, and the idea that this relationship is one of over-dependence upon and an over-preoccupation with the past. The mention of 'chain' suggests a binding relationship where the individual is left unable to escape his past self and past actions. There is thus a sense here of memory

as an obstructive capacity, preventing the individual from fully engaging with present or future projects.

It is particularly easy to understand this thought with the idea that an individual is plagued by painful memories, which act as 'a dark, invisible burden' (UM II:61). For example, we might think of trauma victims who are unable to move on from past events and whose lives continue to be dictated by these events. But this could equally happen with more pleasant memories: we can also imagine someone who is overly nostalgic and unwilling to let go of the past: in this scenario, the memories will have positive affective experiences associated with them and the individual may wish to 're-live' these moments-perhaps to such an extent that they are disengaged with or even hostile to the present and/or the future. In this passage, Nietzsche remains neutral on the actual content of the memories or the affective experiences associated with them: the point seems to be that, regardless of the semantic content of the memories and their associated affective content, we can be overly-dependent on them such that this has a debilitating effect.

Later Nietzsche crystallises this over-preoccupation with the past into a single phrase 'it was', describing it as 'that password which gives conflict, suffering and satiety access to man so as to remind him what his existence fundamentally is – an imperfect tense that can never become a perfect one [*ein nie zu vollendendes Imperfectum*]' (UM II:61). This reference to the imperfect could suggest something bittersweet. As a tense, the imperfect at once situates an action in the past (meaning that this is something that we no longer have access to), whilst suggesting that there is a continuous or repetitive aspect to the action (meaning that there is some sense that the action was never fully completed). In this way, Nietzsche suggests that our actions and projects will inevitably be disturbed or frustrated, and that our memories, whether good or bad, serve as reminders of this fact. Fixating on the past in this way, therefore only makes us aware of transience. Another way of understanding the 'it was' is as continued or repeated reverberation of (painful) memories. It is perhaps for this reason, that in the third *Untimely Meditation*, Nietzsche suggests that our relationship to memory and our past is in fact one of fear:

> everyone is familiar with the strange condition in which unpleasant memories suddenly assert themselves and we then make great efforts, through vehement noise and gestures, to banish them from our minds: but the noise and gestures

> which are going on everywhere reveal that we are all in such a condition all the time, that we live in fear of memory and of turning inward.
>
> (UM III: 158)

An over-preoccupation with one's memories therefore has the potential to be a distracting or even incapacitating force. It seems that with the idea of moving on from x or not dwelling on x, Nietzsche is suggesting that we do not get tied down to the past and avoid this inertia or continuation of suffering.

In *On the Genealogy of Morality*, Nietzsche gives another insight into how one's relationship with the past can be troubling, namely the idea that individuals retain and utilise negative memories to fuel current negative affective experiences. In this work, Nietzsche contrasts the 'prodigious memory' (Deleuze:1962:108) of the slaves to the ability of the nobles to forget past misdemeanours (whether their own or those of others)[32].

Constantly oppressed by the nobles, the slaves amass unpleasant experiences and memories of these experiences, and allow associated negative affective content to linger in their mental economies. In *Ecce Homo* Nietzsche describes what happens if these memories and the associated affective content is retained–individuals are unable to 'get rid of anything, you do not know how to get over anything, you do not know how to push anything back, -everything hurts' ('Why I am so Wise' 6). Without forgetting 'men and things come importunately close, events strike too deep, the memory is a festering wound' ('Why I am so wise' 6). Here the thought seems to be that if one's mental economy becomes oversaturated with (presumably painful) memories, this creates trauma for the individual – a thought echoed in *Thus Spoke Zarathustra* when Nietzsche calls memory the will's 'loneliest misery' where the individual 'impotent against that which has been' is reduced to 'an angry spectator of everything past' (TSZ 'On Redemption').

[32] 'A man like this shakes from him, with one shrug, many worms which would have burrowed into another man' GM II:10.

This seems to be precisely what occurs in the case of the slaves – rather than forgetting painful events they have undergone, they hold on to these moments, building up feelings of *ressentiment*: 'he knows all about keeping quiet, not forgetting, waiting, temporarily humbling and abasing himself' (GM I:10). The danger here is not merely that these memories are retained but rather the individual then utilises them in further self-destructive projects. In GM III:15 Nietzsche suggests that the slaves mine their unpleasant memories they have stored up, to further facilitate feelings of *ressentiment* within their psychology:

> The sufferers...enjoy being mistrustful and dwelling on wrongs and imagined slights: they rummage through the bowels of their past and present for obscure, questionable stories that will allow them to wallow in tortured suspicion, and intoxicate themselves with their own poisonous wickedness.
>
> (GM III:15)

These negative memories provide them with 'evidence' for their suspicions and foster feelings of mistrust. In this way, their memories bolster negative affective content – the slaves use memories to amplify and justify their current emotional states. This activity contributes to a sense of self which is primarily reactive- the slaves constitute themselves in such a way that their present self is determined by resentment towards past experiences. In this passage, with the references to intoxication and poison [*am eignen Gifte der Bosheit sich zu berauschen*] Nietzsche seems to suggest that this process is unhealthy.

We are now in a position to see why Nietzsche advocates forgetting in the sense of moving on from x or not dwelling on x. It permits a healthier relationship with the past: enabling the individual not to feel tied to their past such that it prevents them from engaging in future projects, and allowing them to disassociate from negative affective content.

2.3.1 How to Forget

In what way should we 'move on' from unpleasant memories? There are several options. The 'healthy' form of forgetfulness Nietzsche advocates, is, as Swanton (2015:163) notes, to be distinguished from a) 'an attitude of non-caring or insouciance about harms one has

caused', and b) 'a form of weak forgetfulness which is a form of escape from self'. In this way, Nietzsche does not promote callousness or total indifference. Such an attitude would not involve any real confrontation with past events. In order for painful memories to be fully processed they must first be acknowledged (we may think of the confrontational lion in 'The Three Metamorphoses' here).

The 'strong' form of forgetting Nietzsche has in mind has more to do with moving on and not dwelling on past transgressions via a total *absorption* of these experiences. The failure of an individual to fully process memories in this manner results in self-hate, guilt, and regret and is a sign of weakness. So how are we to understand this notion of absorption?

If we return to GM II:1 we can see how Nietzsche utilises the analogy of digestion to illustrate this idea:

> Forgetting …. is responsible for the fact that what we experience and absorb enters our consciousness as little while we are digesting it (one might call the process "inpsychation") as does the thousandfold process, involved in physical nourishment—so called "incorporation".
>
> (GM II:1)

The analogy here is complex[33] and aligns the mental with the physiological. Nietzsche mentions three processes: (i) absorption, (ii) digestion, and (iii) incorporation. Absorption implies that material is taken in indiscriminately just as a sponge soaks up water. The idea of digestion, however, is more nuanced. The act of ingesting food is usually reflectively conscious but the processing or digesting of the food is not – it merely occurs automatically without our being reflectively conscious of it (unless of course you suffer from indigestion and consciously experience stomach pains). Nietzsche seems to suggest here that the case of experience and memory is analogous – we experience events consciously but the process of what he calls "inpsychation" or "incorporation" is unconscious. If the experiential material is something which seems particularly resistant

[33] My thanks to Manuel Dries for his extremely helpful comments regarding this analogy.

to being incorporated (for example, experiences that generate anger or jealousy or something like a trauma, i.e. things that are difficult to move on from), then this could be experienced consciously. Emotions like anger or jealousy seem particularly salient to us. In such cases, Nietzsche suggests that, as a 'positive faculty of repression' the forgetting capacity would become operational, and inhibit this material from rising to the surface of consciousness.

At this point, we might think that Nietzsche is presenting something like an account of Freudian repression where material is effectively censored. In BGE 68 Nietzsche gives the example of the drive to pride dominating and suppressing memories: '"I have done that", says my memory. "I cannot have done that," says my pride, and remains unyielding. Eventually – memory yields.'[34]

However, Nietzsche is not being proto-Freudian. As Conway suggests what Nietzsche is advocating is a 'deep forgetfulness', i.e. 'a complete metabolisation of experience rather than the repressive forgetfulness that Freud's later concept of the Unconscious introduced into the mental economy, where what is put there clamours to be made conscious and so is not deeply forgotten' (2012: 38). On Freud's model the suppression is only temporary, and material can continue to manifest itself via maladaptive behaviours and anxiety, thus still exercising a potentially toxic effect on the individual. For example, we might think of someone who, as a child, was bullied mercilessly, managed to repress those experiences, but as an adult is often overly defensive at work, and as a result is a very difficult colleague. In this scenario, it seems that the original negative affective valence of that childhood experience still plays some motivational component in the individual's behaviour in later life. This is a crucial difference between

[34] In this passage Nietzsche presents something similar to Freudian repression, and explains it through the activity of the drives: here we see an agonistic process whereby the drive to remember loses out to the drive to pride.

Nietzschean forgetfulness and Freudian repression.[35] In order to understand this difference, we must return to the notion of digestion.

Nietzsche's analogy suggests that just like food, experiential material is not taken in in a haphazard manner. Just as nutrients are extracted from the food and incorporated into the body to serve certain ends (e.g. calcium strengthens bones), experiential material is also sorted and processed. Once the material has been taken in and made use of, it is then properly eliminated with no harmful residual existence: '… forgetting plays a role in the regulatory process that permits us to appropriate our experience such that we take from it what is necessary and rid ourselves of what is not' (Acampora 2006: 149). This seems to be an important point: if forgetting is really meant to be a mechanism to maintain mental health -'a doorkeeper, a preserver of psychic order, repose, and etiquette'- then ensuring that harmful material is properly eliminated seems to be more effective than simply repressing it, whereby it can re-emerge in other (potentially) destructive ways. The use of 'he cannot "have done" with anything' implies that there needs to be this finality of moving on in forgetfulness.[36]

Nietzsche returns to the analogy of digestion later in the *Genealogy*, explicitly aligning the ability to process experiences in this manner with a strong and healthy individual:[37]

> a strong and well-formed man digests his experiences (including deeds and misdeeds) as he digests his meals, even when he has hard lumps to swallow. If he 'cannot cope' with an experience, this sort of indigestion is as much

[35] This is not to say that on the Nietzschean model it is never the case that there is material which is not forgotten – this process of 'total metabolisation' is for Nietzsche an ideal, and only achievable by certain individuals.

[36] It seems that in order to achieve this finality, forgetting must work upon not only occurrent beliefs but also dispositional beliefs.

[37] Cf. UM II: 63: 'There are those who are so little affected by the worst and most dreadful disasters, and even by their own wicked acts …The stronger the innermost roots of a man's nature, the more readily will he be able to assimilate and appropriate the things of the past'; GM I:10: 'of what concern are my parasites to me? It [a society of strong individuals) would be entitled to say. May they live and prosper: I am strong enough to allow that!' This ability to move on enables the individual to remain a 'yea-sayer.'

> physiological as any other- and often, in fact, just one of the consequences of that other.
>
> (GM III:16)

This ability to 'cope' with an experience is translated into the physiological. In this way, we see how forgetting an experience and digesting it is both a sign of and a means to preserving or augmenting a strong constitution.[38]

Processing experiences in this fruitful manner – via assimilation and appropriation- affords strong individuals even more strength and endurance.[39] At this point, we ought to emphasise the distinction between evaluative content of an experience on the one hand, and affective valence of an experience, on the other. The former can be understood as an assessment of the experience, e.g. 'breaking my leg was horrifically painful', 'I didn't enjoy playing in my grade 2 piano exam at the time', 'I was really proud of myself when I passed my driving test', whilst the latter can be seen as relating to the degree and intensity of the emotion experienced at the time, such that the individual is able to, to an extent, re-live the experience.

I suggest that an experience is digested if and only if: a) the evaluative content (positive or negative) of the experience is retained, b) any positive affective valence associated with the experience is retained, and c) negative affective valences associated with the experience are reduced or removed such that they no longer play any motivational role in the individual's attitudes and/or behaviour (unlike in the example of the bullied child). Not everyone is able to process experiences in this manner. So called 'weaker types', with a substandard drive configuration, are more likely to retain these past experiences in full affective detail, i.e. retaining the negative affective valences, which weakens them further: 'there are people who possess so little of this power that they can perish from a single

---

[38] The idea that forgetting is an integral capacity is underlined by Nietzsche's remarks on mnemonics and the memory of the will in GM II:3 where he suggests that it is only through great pain that we can counteract forgetting: 'a thing must be burnt in so that it stays in the memory: only something that continues to *hurt* stays in the memory.'

[39] Endurance is highly prized by Nietzsche: see HAH I:263; TI 'Arrows and Epigrams' 8; WP 910.

experience, from a single painful event, often and especially from a single subtle piece of injustice, like a man bleeding to death from a scratch' (UM II:62).[40]

At this point, we might question whether the idea of retaining negative affective valence need always weaken the individual, as Nietzsche suggests. It seems that this will depend on the specific emotional form this negative affectivity takes, for example there may be cases where the emotion of shame has strengthening effects, such as compelling the individual to cultivate certain character traits, which will ensure he will not be confronted with those sorts of difficult (shame-inducing) situations again. It may be better then, to restrict Nietzsche's claim to self-reflexive emotions that signal situations the individual cannot do anything about (these may well arise in the bullied child case). In such cases, it makes sense to say that the negative affective valence would have a debilitating effect on the individual if retained.

The idea of incorporation (as 'digestion' of experiences) engendering strength gains further traction in GS 110: 'the strength of knowledge lies not in its degree of truth, but in its age, its embeddedness, its character as a condition of life.' The suggestion here is that in strong individuals only memories which serve to strengthen their hierarchy of drives (be these pleasant or unpleasant memories) are permitted to remain. The 'degree of truth' of such a mental item is not of primary importance.

Nietzsche extends the metaphor of the physiological beyond his analogy of forgetting as digestion to include medical language. The emphasis on forgetting as recuperation is striking. In GM I:10 he declares: 'to be unable for any length of time to take his enemies, his accidents, his misdeeds themselves seriously – that is the sign of strong, full natures in which there is an excess of formative, reconstructive, healing power which makes one forget,' and in UM II: 62: 'the capacity to develop out of oneself in one's own way, to

[40] We may recall the slaves of the *Genealogy* who possess a 'prodigious memory' (Deleuze: 1962: 108), which leads to guilt and ressentiment. They must remember precisely so they can continue to deny and 'say no to what is 'outside,' what is 'different,' what is 'not itself.'' (GM I:10). This is not to say however that Nietzsche is wholly negative about memory. In (UM II:63) he suggests that a balance ought to be struck between memory and forgetfulness: 'cheerfulness, the good conscience, the joyful deed, confidence in the future – all of them depend…on one's being just as able to forget at the right time as to remember at the right time.' Again, it seems that only strong types will be able to employ memory in a healthy way.

transform and incorporate[41] into oneself what is past and foreign, to heal wounds…'[42] The image of the self which emerges here is one of malleability: a strong individual heals themselves via this 'plastic power' (UM II:62) of forgetting by incorporating painful experiences in such a way that their damaging effects are curbed, i.e. the negative affective valence of the experience is diminished.[43]

It can, of course, be objected that, by Nietzsche's own standards, remembering one's experiences vividly in fully affective detail is a sign of strength or health. We might think of so-called "artistic types," such as writers, poets, artists, musicians, or actors who often draw on their experiences in their works, re-engaging with and challenging the affective valence of these experiences. One of Nietzsche's favourite examples of an "artistic type" is Goethe, and it seems unlikely that Goethe necessarily forgot full affective detail in processing his experiences. One way we might account for such a worry is to say that in the case of these individuals the drive to artistic creativity, being particularly dominant, is able to channel and transform this affective valence. The idealised state of forgetting I discuss in this chapter is thus not the only ideal for healthy individuals, but rather one amongst other possible ideals which are conducive to flourishing, such as the artistic ideal.

The language of convalescence Nietzsche uses suggests that, through forgetting, the self is able to regulate itself. By absorbing and incorporating experiences in such a way as to avoid the damaging effects, the self is able to stabilise itself and preserve its strength. Forgetting aspects of certain experiences may affect the way other drives and affects

---

[41] Huddleston (2017:13) discusses the merits of incorporation over extirpation, suggesting that for example with elements of the self, such as pride, lust to rule, sensuality etc., 'the healthier response, often, is to be able to incorporate or integrate these elements into the whole that one potentially is.'

[42] Cf. 'Only now do I believe you healed: for healed is who forget' (7. Friedrich Nietzsche, Zwiegespräch, in Gedichte, R. Kray and K. Riha, eds. (Frankfurt: Insel Taschenbuch, 1994), p. 54: "Jetzt erst glaub ich dich genesen: / Denn gesund ist, wer vergass.")

[43] This narrative of resilience to sickness via incorporation is also extended to whole cultures: 'it is precisely at this injured and weakened spot that the whole body is as it were *inoculated* with something new; its strength must, however, be as a whole sufficient to receive this new thing into its blood and to assimilate it… a people that becomes somewhere weak and fragile but is a whole still strong and healthy is capable of absorbing the infection of the new and incorporating it to its own advantage' (HAH I:224)

interact with each other, subduing or augmenting their activity according to the experiences forgotten, affording the individual a sense of therapy.

The phenomenon of self-regulation seems vital for allowing an individual to take up second-order attitudes against their experiences. Janaway invokes this notion of second-order attitudes in relation to taking up an affirmative stance and being 'well-disposed' to one's life, noting that:

> Nietzsche does not speak of assessing or judging the amount of good that a life contains. Rather his question seems to be: given the amount of suffering, lack, boredom, and triviality in a life, how well-disposed can you be towards it?
>
> (2009:196)

Becoming well-disposed involves second-order attitudes. Janaway suggests that:

> numerous events in any life will be undergone, remembered, or anticipated with a negative first-order attitude; but that is compatible with a second-order attitude of acceptance, affirmation, or positive evaluation towards one's having had these negative experiences.
>
> (2007:257-8)

It seems plausible that we might take up such attitudes of acceptance or positive evaluation with regards to unpleasant or difficult scenarios, in which there is some kind of instrumental payoff: consider for example, a painful trip to the dentist – here we can recognise this event as involving a negative first-order attitude, but there seems to be some purpose to it which we would positively evaluate, i.e. maintaining good dental health. This kind of scenario is in keeping with a thought that Nietzsche sometimes seems to express – what doesn't kill me, makes me stronger. But this thought, and Janaway's suggestion, seem somewhat difficult to transpose to extreme cases, e.g. traumatic episodes where there does not seem to be any instrumental payoff available. I suggest that from Nietzsche's remarks on forgetting, we can infer that, at times at least, he advocates something different to the 'what doesn't kill me makes me stronger' stance, and seems to put forward a means of treating our experiences, such that we can eventually arrive at a positive, or at least neutral, second-order attitude.

Forgetting facilitates such second-order attitudes: it allows a process of digestion to take place. We may acknowledge and remember that a past event was unpleasant (thus retaining evaluative content, both positive and negative), but we manage not to dwell on this in such a way that it paralyses us. For example, if every time I ride a bike, I remember the very first time I attempted this, and the painful crash which resulted, I would not be able to cycle without some trepidation. Such a memory may of course be recalled in exact detail (i.e. the evaluative content: 'I crashed into a tree', 'I needed stitches' 'it was a horrible incident' and so on) but the strong affective valence of fear, pain, and humiliation is dialled down, and consequently we will not re-live the experience in the same manner. Recalling and re-living seem to be phenomenologically different episodes, and the distinction is important.[44]

In this way affirmation need not always be an enthusiastic celebration but can instead be characterised as a somewhat neutral acceptance. The unpleasant event does not affect me in the same way anymore (although I am able to acknowledge how and why it was so unpleasant); I am able to move on from it having taken the strengthening 'nutrients' from the experience. Forgetting therefore should not be understood as re-editing or censoring but as this phenomenon of digestion.

The sense of affirmation that forgetting affords is particularly important for two key themes in Nietzsche's practical philosophy, namely Amor Fati (EH 'Why I am so clever' 10; GS 276) and the Eternal Recurrence (GS 285; 341; TSZ 'On the Vision and the Riddle'). In Thus Spoke Zarathustra Nietzsche states that 'the child is innocence and forgetting, a new beginning, a game, a wheel rolling out of itself, a first movement, a sacred yes-saying' (I: Three Metamorphoses). Here forgetting is explicitly linked with affirmation. This is not the place for a full analysis of Amor Fati or the Eternal Recurrence, and, bearing in mind that there is little interpretative consensus, I do not take a strong stance here, but, for the purposes of this chapter, I will refer to one popular interpretation of the themes as existential tests or challenges cf. Ridley 'the best way to construe Amor Fati throughout

[44] We can imagine how a trauma victim would be unable to take up second-order attitudes against their experiences precisely because they continue to *re-live* their experiences, not just remember them.

Nietzsche's work, then, is as an ethical injunction concerning one's attitude towards the world, rather than as a (disguised) metaphysical thesis about how much of the world is necessary' (2005: xvii). Understood this way, it is worth noting the tension that emerges between: 1) the strong rhetoric of 'innocence', and 2) the idea that in order for the test to be real, we have to remember the totality of our life's events.

It seems that, on this picture, not everyone will be able to resolve the tension and pass the test. Those who remember the full affective valence of past events may not pass the test (as the episodes would be deemed too unpleasant to affirm), neither may those who either did not have a full-blown affective experience in the first place and were completely disengaged (e.g. the masters in the *Genealogy* who exhibit a 'water off a duck's back' attitude), or those who have forgotten the affective experience in a traditional sense (i.e. that memory is not available for recall). We could imagine, however, how those who are able to remember past events with the second-order willing which forgetting permits, whereby the affective valence of events has been amended or dialled down, could be successful.

The notion of absorption via forgetting is a key component of a healthy individual for Nietzsche. It also attests to his conception of health as an ideal rather than as a norm. Contrary to a static, non-personal notion of health, Nietzsche instead suggests that health is individual and contextual, and something which must be constructed[45] (GS 120; HAH I:286). Health is dependent upon 'your goal, your horizon, your energies, your impulses, your errors, and above all on the ideals and phantasms of your soul'. The inclusion of 'errors' here is important – it suggests that we have to find some means of dealing with such experiences. Forgetting presents itself as a viable method for processing 'errors' by limiting the affective impact of unpleasant recollections: 'that which such a nature cannot subdue it knows how to forget' (UM II:63). In a strong individual, forgetting will be especially active, treating material in response to other drives (e.g. the drive to pride) such that the individual is able to work towards goals.

[45] Nietzsche therefore seems to align health with Becoming rather than Being.

2.3.2 Mental "Space" and Reinterpretation

The proper 'digestion' of experiences due to forgetting has the added benefit of freeing up mental "space". Such space should not be characterised in terms of storage or capacity; i.e. it is not a matter of deleting memories to make room for new ones. Rather, according to Nietzsche, forgetting amends the affective valence of our memories, and avoids an overloading or congestion of affective states which would impede further attitudes and actions. (On the Freudian model, by contrast, the problem is that for the analysand the affective valence of repressed past events does remain unchanged.) The "space" can thus be understood as a freeing up of mental energy. In this way, the capacity to forget performs a 'deep-clean' procedure within the mental economy:

> To close the doors of consciousness for a time…a little quietness, a little *tabula rasa* of the consciousness, to make room for new things, above all for the nobler functions and functionaries, for regulation, foresight, premeditation ….so that it will be immediately obvious how there could be no happiness, no hope, no pride, *no present*, without forgetfulness.
>
> (GM II:1)

This space is therefore essential for: a) the well-being or 'happiness' of the individual, b) their agency, and c) their ability to engage in the re-interpretative project. Nietzsche first links happiness and forgetting in 'On the uses and disadvantages of history for life' where he suggests that 'it is always the same thing that makes happiness happiness:[46]the ability to forget.' Remembering and re-living all past events in full affective detail has a highly detrimental effect upon an individual, congesting their memory and overloading their affective responses.[47] In the same work, Nietzsche suggests that forgetting is part and parcel of agency – without this phenomenon we would be left unable to act:

---

[46] The happiness Nietzsche describes here is contrasted with 'the dark, invisible burden' and 'pressure' of the past (UM II:61).

[47] This claim strikes a chord with cases of hyperthymesia, a condition whereby individuals possess an abnormally detailed autobiographical memory. One hyperthymestic Jill Price (referred to by the

> Imagine the extremest possible example of a man who did not possess the power of forgetting …. such a man would no longer believe in his own being, would no longer believe in himself…he would in the end hardly dare to raise his finger. Forgetting is essential to action of any kind.[48]
>
> (UM II:62)

Forgetting here is seen as a facilitator or enabler of action and further experiences. By limiting any affective congestion, the individual is able to properly entertain and engage with future-directed thoughts and projects, i.e. the 'nobler functions' of 'regulation, foresight, premeditation'.

Nietzsche further develops the idea that forgetfulness is necessary for action (in that it facilitates other drives),[49] and specifically value creation, in the second *Untimely Meditation* by aligning it with the 'unhistorical'. The 'unhistorical' is characterised as 'an atmosphere within which alone life can germinate, and with the destruction of which it must vanish.'[50] Nietzsche describes the unhistorical as a creative hotbed for values. Initially the unhistorical individual, who is not tied down by thoughts of the past, is blinded and overwhelmed by the present, making it difficult for him to evaluate (UM II:64). However, this state is the necessary precursor to value generation and contains within it all denominations of values: 'it is the condition in which one is the least capable of being just….and yet this condition unhistorical, anti-historical through and through – is the womb not only of the unjust but of every just deed too.' In effect, we must wipe the

---

anonymiser 'AJ') reported her condition to Parker, Cahill, and McGaugh (2006) as being 'totally exhausting' and 'a burden.' Marshall (2011) in examining the case of AJ suggests that 'it isn't just about retaining the significant stuff. Far more important is being able to forget the rest.' This is echoed in Borges' story 'Funes the Memorious' in which Funes suffers due to his incredible ability to recall events. Constant remembering thus has a debilitating effect upon the subject and is inimical to their well-being.

[48] Acampora (2006:159) nicely sums this idea up: 'forgetting, it seems is an important condition for experience – important for giving the shape, form, rhythm, texture, and depth that make the seemingly endless stream of possible objects of concern and attention *an* experience, to recall Dewey's famous distinction, not simply by piling experiences up or onto another, but by taking some away, by encouraging some to fade, recede, fall away.'

[49] 'No painter will paint his picture, no general achieve his victory…without having first desired and striven for it in an unhistorical condition' (UM II:64).

[50] Life is a notoriously difficult term to unpack in Nietzsche's writings but more often than not, it is associated with value creation and flourishing.

slate clean via forgetting before we can create anew: 'he forgets most things so as to do one thing.'

The forgetting that Nietzsche advocates provides the necessary conditions for value generation. It importantly involves confronting past experiences and values before absorbing them in order to clear mental space and arrive at a 'tabula rasa': before the 'new beginnings' of the child in Zarathustra, we have the lion whose task is not 'to create new values… but to create freedom for itself for new creation (I: The Three Metamorphoses). Forgetting appears as a means to creating this freedom, and a necessary condition for the freedom to create values and interpretations (such as second-order attitudes) in particular.

Conclusion

In conclusion, I have suggested that Nietzschean forgetting is best characterised as a capacity. This allows us to bypass any potential paradox yet still account for its active nature. When recruited by the drives and/or affects, the forgetting capacity underpins the selectivity of our perspectives, meaning that our actions are explained and facilitated not only by drives but also by forgetting. In this way, I have attempted to draw attention to aspects of Nietzsche's philosophical psychology other than drives and affects.

I hope to have shown that forgetting is a vital component of a Nietzschean self, and one which plays a key role in Nietzsche's practical philosophy, particularly the idea of re-interpretation. When understood as the generation and reinterpretation of ('healthy') values, this project calls for a self which is able to regulate itself, and process experiences in such a way that the health of the mental economy is not only preserved but bolstered. From Nietzsche's remarks on forgetting, it seems that this mechanism goes far to explain how it is that we are able to create room for new interpretations and process old interpretations in such a way (i.e. via incorporation rather than censorship), that the plasticity and strength of the self is preserved. In this way forgetting allows the self to be in a state in which the conditions for generation of values are optimum.

In a post-Freudian context, it might seem that the idea of forgetting as necessary to a healthy life is widely recognised. So what does Nietzsche contribute to this recognition? The apparent truism of forgetting being required for health is in actual fact ambiguous. I have argued that upon closer inspection it emerges that Nietzsche presents a very different (and more positive) theory of forgetting than that of Freud's theory of repression. The emphasis here is on confrontation and incorporation of experiences, as opposed to mere suppression or avoidance.

As such Nietzsche offers a far more therapeutic understanding of forgetting, and by characterising it as active, one which is not passive or degenerative. This component of the mental economy plays an important role in self-regulation.

3. Memory, First-Person Subjectivity, and Self-Constitution in Nietzsche

Introduction

In the previous chapter, we saw how Nietzsche holds that we can, at times, have an unhealthy relationship to the past, and that memories can sometimes have a detrimental effect on us. In these cases, past experiences have an inordinate and/or negative impact on individuals' emotional lives, attitudes, and actions. This chapter will explore what a healthy relationship with the past would look like according to Nietzsche, and the idea that autobiographical memory in particular can play a role in this healthy form of self-constitution.

I suggest that autobiographical memory can be understood as a plastic and creative capacity, which is responsive to, and works alongside the drives. The interactions between autobiographical memory and the drives work to create a narrative identity for the individual through the reinterpretation and incorporation of events. Whilst memory has often been invoked in relation to discussion of diachronic identity cf. Lockean accounts[51], where identity is understood as numerical identity persisting over time, the notion of narrative identity, which I suggest we find in Nietzsche, is much richer. Narrative identity is more sophisticated in that an individual's narrative coheres with a deeper sense of how events are related to their character (understood on the Nietzschean picture, as their constitution of drives), and/or future projects. We should note that there is an assumption of diachronic identity underlying the reinterpretive activities of the drives and autobiographical memory, as they interact to create an individual's narrative identity. In this

[51] This notion of identity is relevant for legal scenarios, e.g. whether or not I am the same person who robbed a bank two years ago.

chapter, I suggest that this notion of narrative identity is helpful in filling out the kind of 'I' which emerges on a Nietzschean account.

The idea of a narrative self in Nietzsche has been interpreted in a number of ways. Passages like BGE 54 might encourage us to view this narrative self as fictionalist or instrumentalist, cf. Dennett (1992), where the ontological status of the self is akin to something like the centre of gravity, i.e. a 'useful fiction'; cf. Gardner (2009). Meanwhile, passages such as GS 290, could, as Nehamas (1986) suggests, lead us to view the self as a work of art. However, as Gardner points out, adopting a fictionalist view of the self is problematic with regards to Nietzsche's practical philosophy, and Rutherford (forthcoming:7) cautions us against 'giving inordinate weight to an aestheticized conception of the self.' Following Tubert (forthcoming) and Rutherford (forthcoming), I will therefore appeal to a more robust notion of a narrative self as put forward by Schechtman (1996, 2011) where narrative 'is not simply a static set of facts' about an individual 'but rather a dynamic set of principles, a basic orientation through which, with or without conscious awareness, an individual understands himself and his world (Schechtman: 1996: 116).

In this chapter, I use *Ecce Homo*[52] as a case-study of how autobiographical memory contributes to a narrative self and fills in notions of 'I', 'me', and 'mine'. Although *Ecce Homo,* has been seen as an aestheticized instance of self-formation, I will use it solely as an example of autobiographical memory, as it functions broadly in individuals. As Bermudez (2017:182) notes, autobiographical memory is a type of memory which is 'self-specifying': the memories involved here are about the individual and their personal past. In this way, first-person perspective is a key feature of autobiographical memory, and is such that it allows the

[52] Commentators often approach *Ecce Homo* with caution 'even R.J Hollingdale, Nietzsche's excellent and sympathetic biographer, has problems with this book' (Ridley: 2005: ix). I am inclined to follow Ridley's suggestion that many passages in this work can be seen 'as helpful dramatizations of a distinctive strand in Nietzsche's later philosophy, a strand having to do with freedom and self-realisation – with what, in the subtitle to *Ecce Homo*, he calls 'becoming what you are.'' If nothing else, the text offers a case-study of memory and can be seen as an interpretation and evaluation of Nietzsche's past. Ridley suggests there are good reasons for viewing the interpretation and evaluation here as honest: 'the first thing to say is that Nietzsche remains fully committed at this period to values of honesty and the intellectual conscience….'How much truth can a spirit *tolerate*, how much truth is it willing to *risk?* This increasingly became the real measure of value for me'… these are not the words of a witting fantasist, or of one bent on falsifying his past.' (2005: xx)

individual to identify themselves in relation to events, and to see these events as meaningful and contributing to the story of 'who they are'. In *Ecce Homo*, as Rutherford (forthcoming: 6) suggests, it is obvious that Nietzsche explores this very question: 'Nietzsche's writings make clear his preoccupation with the question "Who am I?" His final, ironic autobiography, *Ecce Homo*, carries the subtitle "How One Becomes What One Is.' It seems that Nietzsche attempts to find the answer to this question by drawing upon autobiographical memory; in *Ecce Homo,* Nietzsche offers a retrospective of his works and events in his life, and he envisages this retrospective in terms of narrative: 'and so I will tell myself the story of my life [*und so erzähle ich mir mein Leben*].

The first section explores the idea that a narrative self has 'I' content, and that this is provided by autobiographical memory. The kind of 'I' present in the narrative self goes some way to addressing Gardner's concern that 'the eliminativist-cum-fictionalist' theoretical account of the self which emerges if we take drive-theory accounts seriously, is inconsistent with 'the role Nietzsche accords first-person thinking in practical and axiological contexts' (2009:1). In this way, we may be able to locate some degree of first-person subjectivity in a Nietzschean self.

The second section then examines three features of autobiographical memory which help generate the narrative self, and the 'I', and how these manifest in Nietzsche's writings. Firstly, I discuss the idea that autobiographical memory provides an awareness of a self persisting through time. This feature is necessary for two other facets of autobiographical memory, namely how it enables a) a degree of self-knowledge regarding past perspectives, and b) a sense of first-person perspective which is tied to affectivity.

We will see how this narrative self which emerges via autobiographical memory is important for self-regulation: experiences are interpreted and incorporated into a narrative, leading to a sense of selfhood. Without this, individuals are left feeling unanchored and unstable, something which could impede engagement in future projects. A narrative self is therefore important for agency and well-being, meaning that autobiographical memory emerges as an important regulatory capacity in the Nietzschean self.

3.1 First-Person Subjectivity and The Nietzschean self – The Concern

This section will outline how a narrative self involves 'I' content, why this is desirable when giving an account of a Nietzschean self, and how autobiographical memory contributes towards this 'I' content.

When constructing any theory of the self, an integral component is the matter of first-person subjectivity i.e. a sense of 'I'. We tend to align this 'I' with notions of ownership and identity: 'this is *my* unique and particular perspective on my cognitive processes, my affective experiences, and my actions.' First-person subjectivity is the kind of thing which provides an insight into who we are, or who we take ourselves to be, and how we situate ourselves in the world. It is seen as providing access not only to self-knowledge, but to knowledge of mental states in general. Shoemaker states that 'it is essential for a philosophical understanding of the mental that we appreciate that there is a first-person perspective on it, a distinctive way mental states present themselves to the subjects whose states they are' (Shoemaker 1996:157), and Zahavi highlights the importance of first-person subjectivity for accounts of consciousness in particular, claiming that we 'must take the first-personal or subjective givenness of consciousness seriously, since an important and nonnegligible feature of consciousness is the way in which it is experienced by the subject'[53] (2007:67). Korsgaard highlights the importance of first-person subjectivity from a practical standpoint – we seem to want to attribute actions to a self or a person, as opposed to impulses, i.e. something that happens *to* a person: 'to regard some movement of my mind or my body as *my action*, I must see it as an expression of my self as a whole rather than as a product of some force that is at work *on* me or *in* me' (2009:18). This notion of first-person subjectivity therefore seems crucial to an idea of a substantive self.

One immediate concern with Nietzsche's views on the self, which Gardner raises, is that this first-person perspective is difficult to locate if we take the drive theory account seriously.

[53] Zahavi does raise the concern that focusing on first-person perspective may be problematic in that it could give us a 'slanted view' of what subjectivity amounts to (2007:70). For the purposes of this chapter I will not be addressing this concern, but will merely explore whether there is room for first-person perspective in a Nietzschean self.

Commentators generally agree that drives are evaluative, directional, and responsive to other drives and affects cf. Katsafanas 2005, 2009, 2016, Anderson 2012, and Janaway 2009, and that there is a worry that what goes on at this sub-personal level is not accessible from a first-person perspective. On many accounts, it would seem that drives are solely responsible for agency, and that there is no sense of a unified self above and beyond these subpersonal components. As Guay notes (2006: 218) 'these arguments typically aim to show that a distinctive subject or ''self'' is epiphenomenal, or simply absent, by demonstrating that the subject plays no useful role in explaining either belief or action; the illusoriness of the subject is then inferred from its explanatory impotence'.

In many passages, Nietzsche seems to undermine the idea that we have some kind of privileged first-person perspective to ourselves (D 116), leading to a worry that there is a fundamental lack of transparency regarding our attitudes and actions: 'as long as one sees a point in seeking another explanation, there can be additional conceptual resources to invoke, additional drives to posit, and additional explanations to set forth' (Guay: 2006: 219). This non-transparency of the self is particularly problematic with regards to Nietzsche's practical philosophy.

Gardner suggests that 'the eliminativist-cum-fictionalist' theoretical account of the self which emerges is inconsistent with 'the role Nietzsche accords first-person thinking in practical and axiological contexts' (2009:1), and that there remains a 'puzzle' between '(1) a first-person practical standpoint in which we must take it to be *up to us* what is to be done with and about our drives and (2) a third-person theoretical perspective in which *drives decide* what happens or is done with us' (2015:381). Given that in works such as *Ecce Homo,* Nietzsche seems to write with a clear sense of first-person perspective, this concern is especially pressing.

Particularly when we consider a narrative self, this notion of first-person subjectivity seems paramount. Given that the narrative self speaks to concerns regarding who I am, and, how and why I am the way I am, without appealing to first-person subjectivity, it is not clear that the story I am telling is *my* story as opposed to someone else's. If I feel disassociated from the story, the chances are it will not have much bearing on my sense of identity. Whilst I could certainly recount a simple story consisting of facts about myself linked up in some way e.g. yesterday I missed my train by ten minutes, I had a sandwich for lunch, then I read an

article and so on, unless I identify myself in relation to these events through a certain level of first-person awareness and self-referentiality, the narrative does not have the same force. In this way, the kind of narrative involved here is more sophisticated; by referring to attitudes, thoughts, actions, emotions, all of which are imbued with this first-person awareness, I can view past episodes as meaningful to me in some way and as contributing to my sense of identity. A famous example from Perry (1979) hammers home this point – following a trail of sugar on the supermarket floor and searching for the shopper with the torn sack of sugar is one thing, but realising that the shopper with the torn sack of sugar is actually me is another. This introduction of the 'I' concept changes the force of the episode in my cognitive economy in some way, e.g. a different set of responses may be generated.

Autobiographical memories can provide the kind of sophisticated narrative which carries this first-person awareness: we often recall autobiographical memories as embedded within some kind of narrative arc:

> Episodic memory[54] has been thought to involve re-experiencing events from one's past, thus providing its owner with content by which he or she is able to construct a personal narrative, that is, his or her life stories (see, for example, Eakin 2008; Fivush and Haden 2003; Klein 2001; Klein and Gangi 2010).
>
> (Klein 2012:680)

For example, an individual may think of past events in terms of turning points, or definitive moments so that driving a car for the first time is seen as the first real moment of independence, getting married is seen as a shift in one's identity and so on. In this way, the memories attest to certain personality traits, feelings, attitudes, all of which develop within an overarching narrative, which is taken to explain how and why I am the kind of person I am.

With regards to locating first-person subjectivity in the Nietzschean self, Gardner's treatment of this problem is ultimately aporetic. I suggest, however, that some degree of

[54] 'Episodic' memory may be used interchangeably with 'autobiographical' memory for our purposes.

first-person subjectivity may be salvaged on a Nietzschean account, through the use of autobiographical memory and the way in which it provides 'I' content in a narrative self.

The next section will further explore how certain features of autobiographical memory enable this sense of 'I' present in a narrative self. Specifically, through the way in which autobiographical memory provides: a) an awareness of a self persisting through time, b) an insight into past perspectives i.e. a degree of self-knowledge, and c) a first-person perspective which is tied to affectivity. I suggest that the latter two features are particularly important in allowing a sense of first-person subjectivity, I will draw on examples from Nietzsche's writings to illustrate these three features of autobiographical memory.

3.2.1 A Self Persisting through Time

Sometimes Nietzsche presents an awareness of ourselves as situated in time as troubling. In the second *Untimely Meditation*, he suggests that in gaining an awareness of our temporality, we also gain an awareness of the fact that our projects can be hindered, frustrated or interrupted. We are confronted with the 'it was… an imperfect tense that can never become a perfect one [*ein nie zu vollendendes Imperfectum*]' (UM II:61). This can create feelings of dissatisfaction and disillusionment for the individual – if we come to feel that our endeavours are inevitably futile, then it seems very difficult to retain any motivation or direction vis-à-vis our actions, and we become aware of transience: 'existence is only an uninterrupted has-been' (UM II:61). But there is another way to think of ourselves as situated in time. This section will explore the idea that through autobiographical memory, we gain a sense of self as *persisting* through time, and that this sets up a more positive relationship with the past.

In autobiographical memory, we are able to view ourselves as enduring: although various actions, events and relationships have come and gone, we have remained to carry out new actions and encounter new events and relationships. The main character, as it were, is a constant, though the plotlines change. Moreover, this idea of persistence leaves open the possibility that an individual may develop and improve. Particularly if those past events are remembered as being unpleasant in some way, the fact that the individual also remembers

themselves as having come through those events affords them with a feeling of triumph. Acknowledging this facet of autobiographical memory may provide the individual with a sense of coherence or unity[55] – qua protagonist they have prevailed even in adverse circumstances. This can certainly be seen in passages in *Ecce Homo*, particularly when Nietzsche reflects on his poor health over the years cf. 'Why I am so wise 1' 'Why I am so clever 2.'

In drawing upon these past memories, the individual is able to use their experience of the past as relevant for the present, something which Nietzsche touches on in his discussion of the 'historical man' in the second *Untimely Meditation*:

> Let us call them historical men; looking to the past impels them towards the future and fires their courage to go on living and their hope that what they want will still happen, that happiness lies behind the hill they are advancing towards. These historical men believe that the meaning of existence will come more and more to light in the course of its process, and they glance behind them only so that, from the process so far, they can learn to understand the present and to desire the future more vehemently.
>
> (UM II:65)

Whilst Nietzsche advocates caution with regards to focusing exclusively on the past, he does allow that some anchoring in past memories affords the individual with an awareness of different modes of existence. Linking past selves with the present self, allowing them to inform the present, and assimilating experiences forms a sense of unity with the individual's psychology such that they are better prepared for future actions. In this way, Nietzsche introduces a more refined notion of happiness (which was previously only equated with a capacity to forget), and suggests that we require some balance between acknowledging our past, and not dwelling too extensively upon it:

[55] Cf. Hales and Welshon (2000:168) 'Nietzsche maintains that memory is the lynchpin required for this sense of a unified self that persists through time to arise'

> Cheerfulness, the good conscience, the joyful deed, confidence in the future - all of them depend, in the case of the individual as of a nation, on the existence of a line dividing the bright and discernible from the unilluminable and dark; on one's being just as able to forget at the right time as to remember at the right time; on the possession of a powerful instinct for sensing when it is necessary to feel historically and when unhistorically…the unhistorical and the historical are necessary in equal measure for the health of an individual, of a people and of a culture.
>
> (UM II: 63)

The 'unhistorical' here is understood as an absence of the awareness of temporality where the individual lives completely in the present, with no sense of the past or the future. The 'historical' meanwhile sees the individual relating themselves to the past, sometimes excessively so, such that the past becomes a 'dark, invisible burden' and 'a chain' (UM II:61). Whilst an excess of the historical sense can be problematic psychologically, this does not mean that the unhistorical is the superior mode of existence. Instead a balance between the two must be struck.

One way in which this thought becomes concrete is if we apply it to the notion of selfhood, and consider the case of amnesiacs. As Klein and Nichols note (2012), it is not the case that amnesiacs lack a sense of selfhood per se. Amnesiacs are often unable to retrieve memories of themselves living through certain events, i.e. autobiographical memories, which poses a problem for a memory-based sense of diachronic identity, i.e. a sense that an individual is the same person they were, for example, five years ago. In such a scenario, it seems that this sense of identity and the feeling of persisting through time will not be available to the amnesiac. However, as Klein (2012) argues there still remains some thin sense of self:

> Indeed, even when a person experiences cognitive chaos so extreme that it partitions conscious access to his thoughts, memories, and perceptions into intervals of one second (that is, awareness of current thought and perception fades and is replaced each passing second!), a sense of self-unity and continuity reportedly is felt during each momentary slice of awareness. This held for Storring's 'Patient B' (Storring 1936). The patient under discussion was unable to maintain a continued focus on any internal or external topic, event, or scene

> exceeding his brief limit of continued awareness, yet he had a sense of himself as an existing, continuous entity.
>
> (2012: 681)

Klein and Nichols go on to suggest that the amnesiac has a sense of self beyond bare first-person awareness, i.e. waking up in hospital and thinking 'where am I?'

Instead, they posit that the amnesiac is seen as having some unwavering feeling of personal continuity, which is tied to what they term 'trait self-knowledge'. Trait self-knowledge is a sub-system within semantic memory which 'stores information about one's own personality in the form of trait generalizations (for example, Self: Usually stubborn)' (2012:681). This kind of self-knowledge emerges as especially resilient:

> An extensive review of the available research (Klein and Lax 2010) shows that knowledge of one's traits (a) is immune to loss in the face of multiple, often severe, neurological and cognitive insult (including total retrograde and anterograde amnesia, autism, Alzheimer's Dementia, and Prosopagnosia), (b) is empirically dissociable from trait knowledge of others (even well-known others such as the patient's family members) as well as from purely factual, non-dispositional self-knowledge, and (c) can serve as a firmly entrenched foundation for one's sense that one 'is, was, and will be' (for a similar view, see Rathbone et al. 2009).
>
> (2012:682)

In this way, knowing something about one's character traits, e.g. stubbornness (something that might become evident in the present moment e.g. the amnesiac's interactions with the doctors might provide an arena for his stubbornness to manifest), may provide a basis for one's sense of being a continuing, experiencing self even if there is a deficit of autobiographical memory, and one is excluded from one's personal past.

This kind of self-knowledge about one's traits, like whether one is stubborn or not is not sufficient, however, for a narrative self. There is reason to believe that individuals (consciously and/or unconsciously), strive for a thicker notion of selfhood such as can be found in a narrative self. This can be seen from the fact that amnesiacs often confabulate, in

order to supplement their sense of selfhood.[56] As Hales and Welshon note, an individual who is missing autobiographical memories struggles with issues of selfhood: 'no recalled ambitions, episodes, or even an existential context out of which to construct a life, and yet the sense of self must be maintained' (2000:164). In these cases, confabulations act as a compensatory mechanism indicating that individuals, at least to a certain extent, require some anchoring in the past, in order to make sense of their identity. Without it their well-being, sense of direction, and sense of self can suffer. Through confabulating and 'filling in gaps in the story', individuals are able to create a narrative self. This kind of activity is reminiscent of Nietzsche's remarks in the *Genealogy* on the importance of creating meaning and purpose for ourselves and the experiences we encounter: 'any meaning at all is better than no meaning at all' (GM III:28).

It is in this way that Nietzsche's suggestion that a balance between the 'unhistorical' and 'historical' being essential for healthy individuals, becomes pertinent. As Klein, Loftus, and Kihlstrom (2002) highlight, a worrying consequence of an absence of autobiographical memory is the inability of individuals to imagine themselves into the future. The thought here could be that a sense of persistence provides the individual with the reassurance and security needed to feel able to engage in future projects- if one thinks of oneself as entirely transient, it may be difficult to maintain any sense of 'I', 'me', or 'mine', and to find motivations for future endeavours. In providing that consciousness of ourselves as situated and persisting in time, autobiographical memory in particular, has an important role to play creating a sense of stability and allowing for a sense of first-person subjectivity.

The next section will explore how once an individual has this feeling of persisting through time, they are then able to engage in reinterpretative projects related to perspectives and affects.

[56] Confabulation almost always occurs in autobiographical memory.

3.2.2 Self-Knowledge and Reinterpretation of Past Perspectives

Another way in which we might cement our sense of first-person subjectivity and narrative self is by coming to learn about our previous perspectives and attitudes. In *Ecce Homo*, Nietzsche seems to utilise autobiographical memory in order to evaluate past events and to discover what they might indicate with regards to his character. For Nietzsche, the idea that we might acquire self-knowledge is complex. In certain passages, this possibility of self-knowledge hardly seems open at all: 'all actions are essentially unknown' (D116), 'we are necessarily strangers to ourselves, we do not comprehend ourselves' (GM, Preface 1), and yet we must square such remarks with his preoccupation with matters of personal identity and coming to know himself in *Ecce Homo.* In the Preface, Nietzsche declares that 'it seems imperative to say *who I am*' and in the final chapter, the question 'have I been understood?' is repeated four times – we get the feeling that this question is aimed not just at the reader, but at Nietzsche himself.

Whilst the possibility of self-knowledge as pertaining to current actions and attitudes seems difficult on a drive-theory account where 'neither further acts of introspection nor reliance on conscience will be particularly helpful in attaining comprehensive self-knowledge' (Katsafanas:2015:119), I concur with Katsafanas' suggestion that Nietzsche does allow for a degree of self-knowledge to be gained through retrospective evaluation, and by 'looking away from oneself.'

The idea of gaining self-knowledge through hindsight is a familiar one, and one which Nietzsche picks up on in *Human, all too Human*:

> [I]mmediate observation [*unmittelbares Selbstbeobachtung*] is not nearly sufficient for us to know ourselves; we require history [*Geschichte*] since the past flows inside us in a hundred waves; we ourselves are, indeed, nothing but that which at every moment we sense of this continued flowing [*Fortströmen*]….
>
> (HAH II/1 223)

Katsafanas criticises the following interpretation of this passage:

> [T]hat motives or forces present in the deep recesses of the past still persist in the present. For example, the ascetic priests embraced the values of humility, inoffensiveness, and compassion out of ressentiment; so, somehow, when I embrace those values today, I am still doing so out of ressentiment

and argues that this ahistorical claim is highly implausible:

> the fact that one type of person in a completely distinct social, cultural, and historical setting embraced a value because it experienced a certain affect does not entail that I, in circumstances that could hardly be more different, must embrace the same value out of the same affect.
>
> (2015:125)

Instead, Katsafanas argues, Nietzsche's claim in *Human, all too Human* should be interpreted as appealing to the notion of genealogy, i.e. examinations of 'long stretches of human behaviour rather than isolate moments of choice' which allow us to identify: a) gradual changes in and subtle influences upon our behaviour, and b) 'unnoticed aspects of our conceptual schemes through which we experience and interpret the world'[57] (2015: 126).

Although Katsafanas does briefly note the possibility that this kind of genealogical exploration could be applied to one's personal past: 'one way in which I can learn something about myself by looking away from myself is by detecting—either in historical characters, or even in my own history—gradual shapings of behaviour that result from apparently minor factors' (2015:126), he does not take up this suggestion in detail.

It seems, however, that with *Ecce Homo*, Nietzsche gives us a case-study of just that: a retrospective analysis of his own personal past. We should note that this kind of retrospective analysis need not be limited to a genealogy, but can also be achieved via narrative. The important factor is that there is enough evaluative distance. In this case, the distance is twofold - the fact that the objects being analysed are autobiographical memories

[57]Katsafanas (2015:127) gives the example of Nietzsche demonstrating in the *Genealogy,* how different conceptions of agency supplant one another throughout history, and the ways in which this affects an individual's experience of action.

(temporal distance), combined with the fact that this analysis comes in the form of a narrative (narrative distance), creates the requisite distance for 'looking away from oneself'[58] cf. Goldie:

> [T]his way of thinking of oneself, in one way thinking of oneself as another, in another way thinking of that other absolutely as riveted to oneself, is by no means a confusion: it is at the heart of how we think back ironically on our past, and how we make plans and resolutions for our future in the light of our past.
>
> (2012b: 1065)

Nietzsche qua narrator is therefore able to take up an evaluative stance on his own personal history. This kind of evaluative stance is important for the individual, in that it enables them to probe deeper into their narrative: just as we would try to analyse a character in a book in order understand them better, we might endeavour to amass self-knowledge.

In *Ecce Homo,* Nietzsche makes it clear that this retrospective analysis and the answer to the question 'Who am I?' will be achieved through the use of autobiographical memory: 'I looked backwards, I looked out, I have never seen so many things that were so good, all at the same time' [*Ich sah rückwärts, ich sah hinaus, ich sah nie so viel und so gute Dinge auf einmal].* The use of '*sehen*' here is important, having connotations of 'witnessing' or 'beholding' rather than 'searching' – in this way we get the sense that Nietzsche is observing past selves in order to draw conclusions about his character based on past events and choices. As well as narrating the story of his life, he also seems to be, at the same time, 'reading' it. We can therefore see that 'autobiographical memory proper is something more than pure episode recall, where the former is characterised as going beyond 'recalling the who, what, where, and when of an event, to include memory of how this event occurred as it did, what it means, and why it is important' (Fivush et al 2011:322). The way in which these events are narrated can be seen to determine such factors.

In *Ecce Homo*, we see evidence of the evaluation process:

---

[58] 'Psychologists need to stop looking at themselves if they want to see anything at all.' (TI 'Arrows and Epigrams' 35).

> My proof for this, among other things, that I have always instinctively chosen the *correct* remedy for bad states; whilst complete decadents always chose the means that hurt themselves. As *summa summarum* I was healthy; as a niche, as a speciality, I was decadent.
>
> (EH 'Why I am so wise' 2)

Here we see a retrospective assessment of both strengths and weaknesses- Nietzsche acknowledges that he has, at times and in certain ways, been decadent. The passive, almost clinical tone that is evident in this passage, further reinforces the idea that an evaluation is taking place: ''the way I compelled myself no longer to let myself be cared for, served, doctored – this betrayed an unconditional certainty of instinct as to what at that time was needful above all else' ('Why I am so wise: 2). The fact that Nietzsche not only recalls, but also appraises past events through autobiographical memory provides him with an insight into his past actions and attitudes.

The narrative which autobiographical memory creates facilitates this activity of evaluation. As Watson (2015: 270) suggests there is the initial evaluation of an event which takes place at the time – broadly speaking this is composed of the individual's emotions and cognitions. The narrative component of autobiographical memory then:

> adds a second temporal level to the event sequence by transporting it from the past of experiencing the events to the present situation in which it is remembered and narrated. Because of the temporal distance, narrators and listeners can evaluate events retrospectively, externally. These evaluations may differ from the internal evaluations made at the time of the events.

in this way, the individual qua narrator looks back on themselves as a protagonist- the combination of the original experiential content found in the narrative, and the external adjudication from the narrator provides a new perspective on past events.

It might be objected that leaving aside Nietzsche's concerns about the possibility of self-knowledge, there is the further problem of whether self-knowledge is even desirable? There are passages in *Ecce Homo* where Nietzsche appears to warn against attempts to acquire self-knowledge. He declares that '*nosce te ipsum* (know thyself) is the recipe for decline', that there is 'nothing more dangerous than catching sight of yourself with this task' and states

that 'becoming what you are presupposes that you do not have the slightest idea *what* you are.' How are we to square such statements with his clear preoccupation (in the very same work) with uncovering who he is?

A possible resolution might be to say that these warnings are aimed against trying to know your *current* self, and that they do not need to preclude knowledge of perspectives on what you were or have been in the past. In fact, acquiring perspectives on past selves can be fruitful. In *Ecce Homo,* Nietzsche explicitly invokes this idea of perspectives, and suggests that reflecting on and analysing the past throughout his life has allowed him to gather new perspectives:

> To be able to look out from the optic of sickness towards *healthier* concepts and values, and again the other way around, to look down from the fullness and self-assurance of the *rich* life into the secret work of the instinct of decadence – that was my longest training, my genuine experience, if I became the master of anything, it was this. I have a hand for switching *perspectives*: the first reason why a 'revaluation of values' is even possible, perhaps for me alone. –
>
> ('Why I am so Wise' 1)

In works like *On the Genealogy of Morality,* Nietzsche encourages a multiplicity of perspectives, declaring that:

> There is only a perspectival seeing, only a perspectival 'knowing'; the more affects we are able to put into words about a thing, the more eyes, various eyes we are able to use for the same thing, the more complete will be our 'concept' of the thing, our 'objectivity'.
>
> (GM III:12)

In this way, he extols 'the capacity to have one's Pro and Contra *in one's power,* and to shift them in and out, so that one knows how to make precisely the *difference* in perspectives and affective interpretations useful for knowledge' (GM III:12). The ability to amass and manipulate perspectives, which autobiographical memory provides, seems invaluable, not only when applied to Nietzsche's project of the 'revaluation of values', but also to his project of establishing his sense of narrative identity. These perspectives give us an insight into parts of our character that may not have been obvious at the time: 'precisely those little

experiences and internal processes, which we think have been overlooked, in their totality [*Gesamtheit*] depict the individual character most clearly' (BAW 2, 269-272). Similar to how the genealogical methodology uncovered hidden insights vis-à-vis our moral concepts, the kind of retrospective analysis that is possible through autobiographical memory provides a degree of self-knowledge, and enables the individual to take up a critical distance against themselves, in evaluating these past perspectives.

Moreover, it seems that, due to the operations of autobiographical memory, rather than being mere reportage of past events, these perspectives will instead function as *reinterpretations* of these events. Certain aspects of an event may be augmented or diminished in autobiographical memory, such that the perspectives which emerge, will avoid the realism which something like '*nosce te ipsum'* might calls for.

How does this alteration of memories occur? Commentators such as Hutto (2017), Brockmeier (2015), and Schechtman (1994, 2011) suggest that in order to create an overarching narrative of the individual's life, autobiographical memories are often summarised, condensed, interpreted, or even constructed. When it comes to memories relating to our personal past, we are not dealing with simple, discrete entities which have been filed away and can be recalled at will. Instead mental reconstruction is called for, meaning that memories are susceptible to a certain degree of distortion[59]. One question which arises is what exactly determines this summarising, condensing, interpreting, and/or constructing?

In BGE 68 Nietzsche gestures towards a potential explanation. He gives an example of the memory capacity being responsive to, and working in tandem with an individual's drives: ''I have done that,' says my memory. 'I cannot have done that,' - says my pride, and remains adamant. At last memory yields.' In this way, memory can be influenced and shaped by the activity of drives. As Jensen suggests (2013), certain drives will be more relevant and influential than others:

[59] As Jensen notes (2013:197) memory is 'the only faculty by which we can come to a meaningful portrait of ourselves and yet it is manifestly unreliable.'

> Some drive-based mental states like the feelings of pride, regret, guilt, vanity, and nostalgia are intrinsically backward-looking. They bring to mind certain facets of our pasts in accordance with what they require to fulfil their individuated power aims. A drive to pride will call forth particularly self-actuating episodes, while a drive towards regret will make our consciousness attend to a tragic event for whose outcome we were particularly blameworthy.
>
> (2013:197)

In this way, it seems that certain 'backward-looking' drives will dictate if and how a specific memory is altered in some way, and will play a role in shaping the individual's narrative. This kind of resulting narrative will vary according to the 'health' of the individual - placing accents upon and exaggerating certain memories and downplaying others means that, in a healthy individual, drives and memory interact in such a way that the individual's self-esteem is safeguarded and bolstered, and a celebratory, triumphant narrative emerges.[60] For example, when remembering a school Sports Day, I might recall the memory in such a way that the fact that my team won stands out as the defining feature of the memory, rather than the fact that I was picked last for my team. We can imagine that in an unhealthy individual, the interactions between drives and memories might create a very different narrative (or indeed no obvious narrative at all). In *Ecce Homo*, Nietzsche seems to suggest that subpersonal mechanisms can function to ensure that even the unpleasant episodes in one's life can be seen as meaningful and necessary:

> If you look at it this way, even life's *mistakes* have their own meaning and value, the occasional side roads and wrong turns, the delays, the 'modesties,' the seriousness wasted on tasks that lie beyond *the* task […] - - in the meantime, the organising, governing 'idea' keeps growing deep inside – *it* starts commanding [*sie beginnt zu befehlen*], *it* slowly leads *back* from out of the side

[60] Cf. Hutto (2017:200) 'Engaging in narrative practices, those that involve giving accounts in rich storied content, has been shown to correlate positively with mental health. A number of findings demonstrate that people 'who are able to narrate the emotional events of their lives in more self-reflective ways show better physical and psychological health' (Fivush et al 2010:46; see also Fivush et al.2003)

> roads and wrong turns, *it* gets the *individual* qualities and virtues ready, since at some point these will prove indispensable as means to the whole.
>
> ('Why I am so Clever' 9)

As Ridley notes, passages like this have often led commentators to view *Ecce Homo* as an opportunistic way of whitewashing or twisting unpleasant past episodes cf. Jensen (2013:184) 'through selecting, exaggerating, underplaying, and manipulating the raw data of the past, Nietzsche was allegedly trying to forge for himself a palatable self-image for the sake of fate-affirmation…choosing only those colorations of events that one is already happy to affirm.' It is not obvious, however, that this needs to be particularly problematic; as Ridley suggests (2005: xxi), it is not 'very troubling to think that it might be taken as a self-help manual, as a promoter of positive thinking. Positive thinking is surely better than the reverse' and that:

> Confronted with some grim facts about our past, we can of course try to forget it; indeed Nietzsche speaks warmly and often about the value of forgetting. But if that is not possible, it is scarcely opportunistic to try to see it instead as something "that must not be missing," that has 'a profound significance and use precisely for *us*.' To refuse to recuperate what we can out of life is to turn our backs on it.
>
> (2005: xxi)

In this way, *Ecce Homo* emerges as a good example of how the self can recuperate and regulate itself through this use of autobiographical memory. The idea that the drives interact with autobiographical memory, and create a favourable narrative looks very much like a method of reinterpreting and incorporating past events. It seems that this kind of activity is better viewed as incorporation – a phenomenon that Nietzsche repeatedly advocates (GS 110, UM II:62; see Chapter Five for further discussion), rather than whitewashing. It is not so much that past events are glossed over, covered up or that the individual refuses to engage with them, but instead they are utilised as part of a narrative (it could even be the case that the interactions between the drives and autobiographical memory could be so fine-grained that specific elements of unpleasant events may still be retained and occupy a role in the narrative). The individual thus still has to engage with this material; it is not simply put to one side. In this way, the material is not repressed, but rather reinterpreted.

The operations of autobiographical memory, namely the condensing, summarising, interpreting, and constructing can all serve to mitigate unpleasant past events or to provide coherence to random happenings[61]. As Schechtman (1994) suggests this may be a response to the overwhelming amount of information we receive about our lives, which necessitates some sort of processing if we are to garner any valuable knowledge about our lives:

> Our knowledge of what we have done is, however, at the same time knowledge of who we are and what we are like. Memory of what we have done and felt and experienced is one of the most important sources of self-knowledge we have. If, therefore, we wish to glean useful information about ourselves from what we know about our pasts, it will again be necessary to process and interpret the information we take in…The goal of making sense of the unfolding of our lives—of writing our own biographies—would require that we interpret and reconstruct our experiences to create a coherent life history. Anomalous events may thus be recast, representative ones emphasized, and other changes undertaken to make one's past more smooth and comprehensible.
>
> (1994:11)

The plasticity of autobiographical memory and how it interacts with drives, therefore seems crucial to this project of meaning and coherence vis-à-vis personal identity. In this way, as Conway suggests (2004) autobiographical memory recall is steered by two forces (i) correspondence i.e. ensuring the memory is an accurate representation of past events, and (ii) coherence i.e. ensuring that the narrative which develops as a result of these memories is in line with the individual's well-being, and promotes the self in the best possible way. It seems that this force for coherence will be particularly active in healthy individuals, and may, at times usurp a need for correspondence - accurate representations of past-events may therefore be allowed to fall by the wayside in favour of mental well-being and self-regulation cf. Schechtman (1994:9): 'in memory we often condense experiences, presenting

[61] As Goldie notes (2012b: 1065) this notion of 'coherence' does not necessarily require any deep narrative coherence: 'stuff happens, and not all narrative are narrative explanations, which succeed in explaining why stuff happens' but that there is 'coherence in your autobiographical narrative in the sense that it is a narrative of your life.' In *Ecce Homo*, it does look like we get a slightly thicker notion of narrative than this – in evaluating past experiences, Nietzsche does seem to attribute explanations to events.

to ourselves a fictitious event or experience which is a composite of the essential features of a series of real ones.'

Self-knowledge gained through this retrospective analysis and the ensuing narrative has the potential to influence future outlooks and actions. As Wilson (2003) notes there is a bi-directional link between autobiographical memory and identity:

> [I]ndividuals' current self-views, beliefs, and goals influence their recollections and appraisals of former selves. In turn, people's current self-views are influenced by what they remember about their personal past, as well as how they recall earlier selves and episodes. People's reconstructed evaluations of memories, their perceived distance from past experiences, and the point of view of their recollections have implications for how the past affects the present.

In this way, retrospective self-knowledge bolsters a sense of first-person subjectivity in that it provides insight into our past commitments and generates additional perspectives. In healthy individuals, the narrative created by the interaction between the drives and autobiographical memory will be such that it informs an individual's sense of identity, and allows them to make sense of past events without being overwhelmed by them.

3.2.3 First-Person Perspective: Affectivity and Autonoetic-Consciousness

So far, I have suggested that autobiographical memory and the narrative it generates can contribute to first-person subjectivity by providing an awareness of the self persisting through time, and allowing for retrospective self-knowledge. But more needs to be said about the precise nature of first-person involved here. I suggest that autobiographical memory provides us with an 'I' that is tied to affectivity, and that this element of affectivity is particularly crucial for the sense of 'mineness' which is characteristic of autobiographical memory.

Bermudez (2017:7) describes autobiographical memory as crucially involving first-person content and suggests that there are: (i) episodic memory images which preserve an individual's original experience of the remembered event and, (ii) accompanying memory

judgements which have first-person contents, i.e. contents that would standardly be expressed using the first-person pronoun[62]. These judgements originate in and are grounded by the episodic memory images. Often these judgements can involve affective components.

One feature of autobiographical memory which allows individuals to re-engage with the affective content of memories is what commentators often refer to as autonoetic consciousness. This is the idea that individuals mentally 'time-travel' to revisit autobiographical memories and re-engage with them (Tulving 1983, Schacter, et al., 2007; Suddendorf and Corballis, 2007). So, for example, rather than recalling events about one's life as mere facts, e.g. 'I won a prize in a piano competition when I was seventeen', 'I crashed my car into a tree when I was twenty', auto-noetic consciousness would involve an element of mental reconstruction, such that the individual would, in some way, re-live the emotions associated with the event (albeit in a weakened sense: it is highly unlikely that one would experience the exact same affective experience, or experience it to the same degree as originally felt).[63]

With the example of the piano competition, I might re-engage with the anxiety I felt whilst waiting to hear the announcement of the winner at the piano competition, the sense of preoccupation as I was trying to evaluate how well I'd played the pieces, the surprise and elation when my name was announced and so on. This kind of 'mental time-travel' seems plausible if we think of past events which were highly emotionally charged. Whilst memories of facts might just lead to an 'awareness of the past that is limited to feelings of familiarity or knowing' (Gardiner 2001: 1351), the idea of auto-noetic consciousness in

[62] Whilst the majority of autobiographical memories are recalled from a first-person perspective, there are cases where memories can be retrieved from a third-person perspective i.e. where subjects see themselves from the perspective of an external observer (Sutin and Robbins 2008; Rice 2009). As Sutin and Robbins (2008) note in some cases, third-person perspective functions to distance the subject from an event which might seem incongruent with the current self e.g. traumatic events where the subject is trying to avoid psychological pain (McIsaac & Eich, 2004; Kenny & Bryant, 2006) or physical pain (McNamara, Benson, McGeeney, Brown, & Albert, 2005).

[63] We can, of course, think of counter-examples e.g. trauma victims who might well feel the emotions associated with the original trauma to the same degree.

autobiographical memory provides the individual with a certain phenomenological richness. In going back to that moment, I am able to re-engage with the emotions at the time.

Moreover, since I am narrating these autobiographical memories, qua narrator, I may acquire access to additional emotional levels. As Goldie (2009) suggests there can be two fundamental kinds of emotional import and emotional response regarding narratives: 1) an internal aspect i.e. affective content that is integral to narrative, and 2) an external aspect i.e. the response *to* the narrative itself. Goldie highlights how dramatic irony may play a role here: 'the emotion that is internal to the narrative differs in type from the external emotion because the narrator was not aware of certain crucial facts about what happened until later' (2009:99), meaning that:

> you the thinker, the 'external narrator', can think of yourself in past episodes, doing and saying things, and you now do this in a way that enables you to conceive of the episode as having an emotional import that you did not recognize at the time, and thus you are now able to have an emotional response that you did not have at the time.
>
> (2009:103)

Do we see anything like this in *Ecce Homo*? I suggest that Nietzsche describes this kind of mental time travel and re-engagement, and this is particularly evident when he is remembering and assessing his works. Throughout *Ecce Homo*, whenever Nietzsche recounts certain events and moments in his life, he does so not in terms of mere semantic memory, i.e. memories of dates, places, times, but rather by utilising autobiographical memory and autonoetic consciousness to re-construct these episodes and re-live them. This is suggested by the level of sensory detail that appears in the descriptions, for example when he reflects on writing *The Birth of Tragedy*:

> [Y]ou would never believe that it was *begun* in the thunder of the battle of Wörth. I thought these problems out in front of the walls of Metz during the cold September nights when I was serving as a medical orderly.
>
> (EH 'The Birth of Tragedy' 1)

or the way in which, when discussing *Human, all too Human,* he makes his emotions at the time evident, and re-engages with them. First his sense of estrangement and disassociation:

> The beginnings of this book belong in the middle of the first Bayreuth festival; it presupposes a deep sense of alienation from everything around me there…just like a dream…and where was I? I did not recognise anything, I hardly recognised Wagner.
>
> (EH 'Human, all too Human' 2)

Then his frustration

> Ten years had gone by, and during that time the *nourishment* of my soul had come to a complete standstill, I had not learnt anything useful, and I had forgotten an absurd moment for the sake of some scrap of dusty, scholarly junk. To creep through ancient metrists with diligence and bad eyes – that is what I had come to!
>
> (EH 'Human, all too Human' 3)

When starting to discuss *Daybreak*, his use of the present tense conveys a sense of immediacy and implies that he is transported back to that moment: 'my campaign against *morality* begins with this book' [*Mit diesem Buche beginnt mein Feldzug, gegen die Moral*] (EH 'Daybreak' 1) This is presented as a turning point: in this way we can see how narratives are marked out by such moments. Later, we get the sense that this moment still reverberates:

> Even now, when I chance to light on this book, every sentence becomes for me a spike with which. I again draw something incomparable out of the depths: its entire skin trembles with tender shudders of recollection.
>
> (EH 'Daybreak' 1)

In this way, these descriptions suggest that in recalling these memories of events which he experienced first-hand, Nietzsche is able to re-engage with past thoughts, and feelings through this notion of mental time-travel. The memories are not recalled as mere facts about an event but crucially involve a first-person perspective which is often linked to past emotions: this suggests a phenomenological proximity and plausibly contributes to a sense of memory ownership.

It is noteworthy that Nietzsche reflects upon these thoughts and feelings as *his,* and not as the effects of some psychological complex. I suggest that this gives us an insight into the kind

of 'I' which emerges in a narrative self. Riccardi notes that we are accustomed to thinking of first-person perspective as something that is causally efficacious cf. Campbell 'our pattern of use of the first-person is heavily invested in the idea that the self is causally significant' (2012:373). The general consensus is that Nietzsche is deeply suspicious of such an assumption: 'in the passage from BGE 16…Nietzsche stresses how expressions like 'I think' suggest that the mental verb in question – in the example 'think' designates a kind of 'activity' which the 'I' is supposed to be causing' (2015:44). However, it is not obvious that to do away with this understanding of an 'I' as causally efficacious is to do away with the 'I' per se. The use of autobiographical memory here and autonoetic consciousness which I have flagged above, could be seen to gesture towards an 'I' which is merely something to which thoughts, drives, affects and so on pertain.[64] The first-person perspective which is tied to affectivity in autobiographical memory can serve to create a sense of 'mineness.' It could be that this sense of 'I' is enough for a sense of personal identity or narrative self.

Conclusion

In conclusion, I have argued that autobiographical memory as seen in *Ecce Homo* emerges as an important factor in constituting a narrative self for Nietzsche, and goes some way to allowing us to locate the sense of first-person subjectivity required for such a self. I have attempted to begin to address Gardner's concern that the philosophical psychology Nietzsche provides us with does not leave any room for first-person subjectivity: by involving other capacities in the self such as autobiographical memory, alongside the drives, it seems we can retrieve some sense of this.

Autobiographical memory can be understood as a plastic and creative capacity that is responsive to, and works alongside drives. Through the activities of condensing and

[64] We might compare this with the view Anderson (2012) presents of a Nietzschean self as a 'repository.' On this view, the claim that drives and affects are causally responsible for actions and attitudes is maintained, but this does not prevent there being a self which is a distinct psychological object, which stands over and above the drives.

summarising, (in healthy individuals) a coherent and intelligible (and more or less, accurate) narrative emerges. This narrative is such that it provides the individual with a sense of 'I', 'me', and 'mine', something which autobiographical memory achieves through the following ways: 1) by generating an awareness of a self persisting through time, 2) by providing an insight into past perspectives and, 3) by allowing for a first-person perspective which is tied to affectivity. The sense of persistence allows the individual to have a thicker notion of 'I' and to avoid the problems caused by conceiving of oneself as transient, i.e. if someone feels that they have persisted throughout events and will continue to do so, they are more likely to feel motivated to pursue future projects. A degree of self-knowledge is made available to the individual retrospectively, thus allowing them to appraise past perspectives and providing them with some insight into their past characters. In healthy individuals, through the activity of the drives, this evaluative process allows for reinterpretation of past events such that they may be incorporated into the narrative self (and we should note that this process is not to be confused with whitewashing of the past). Finally, autobiographical memory involves a first-person perspective which is tied to affectivity - this allows the individual to re-engage with past affective content and to take up an evaluative stance against such episodes in order to uncover aspects of their emotional life which were perhaps not obvious at the time. In this way, autobiographical memory plays a role in unpacking the phenomenological richness of a self. The combination of the above seems crucial for a narrative self – in order to view this narrative as answering to questions of who I am, and why I am that way, it needs to include a first-person awareness of one's past thoughts, attitudes, actions, and emotions. This sense of selfhood is important for ensuring that the individual feels stable enough to engage in future projects.

4. Self-deception and Pretense in *On the Genealogy of Morality*

Introduction

This chapter examines the use of self-deception as a mode of self-regulation. In *On the Genealogy of Morality* Nietzsche provides us with a case study of self-deception, presenting it as a phenomenon exhibited by the weak and impotent as a response to their oppression: 'the counterfeiting and self-deception of powerlessness' (GM I:13). In the *Genealogy,* this activity of self-deception is contrasted with the strong, healthy capacity of forgetting, which is used by the nobles to avoid occurrences of *ressentiment* (GM II:1), and the implication is that self-deception is the main method of self-regulation open to 'the unfortunate, the downtrodden, the broken…the weakest' or 'the sick' (GM III:14).

In exploring the mental states involved in self-deception, and how self-deceptive beliefs can be maintained, I will focus on the slaves rather than the priest in the *Genealogy*. The reason for this is that the self-deception of the slaves provides us with a more clear-cut and plausible example of self-deception as self-regulation. Nietzsche presents the priest, qua orchestrator of the slave-revolt in morality[65], as power-hungry (a trait displayed by the nobles, causing us to recall that the priestly caste is a faction of the nobles cf. GM I:6-7), and somewhat self-aware – it seems likely that the priest knows that he is using the newly minted values as tools for revenge. The slaves' self-deception, by contrast, appears to have a therapeutic motivation cf. GM III:14, and as such is closer to a more general case of self-regulation: 'that sublime self-deception whereby the majority of the dying, the weak and the oppressed of every kind could construe weakness itself as freedom, and their particular mode of existence as an accomplishment' (GM I:13). In this way, it seems that the slaves' engagement with self-deception is a means of allowing them to feel better about themselves and their situation.

[65] The slave revolt in morality i.e. the introduction of 'slavish' values such as humility, patience, compassion and so on which supersede the previously dominant 'noble' values such as strength, power, pride.

How precisely is this mechanism of self-deception to be understood? Which instances of self-deception occur here? And how is it that this self-deception appears to be relatively stable? Nietzsche presents this phenomenon of self-deception as a key driving force in the *Genealogy,* and it does not appear that the slaves relinquish their self-deceptive beliefs.

In this chapter, I will offer a taxonomy of the self-deceptive beliefs present in the *Genealogy,* drawing a distinction between the slaves' value-facing beliefs ('I believe that humility is intrinsically good'), and their person-facing beliefs ('I am humble'), and exploring how these beliefs can be maintained. I suggest that the biggest potential threat to the slaves' self-deception is their affective experiences, namely their 'secretly smouldering emotions of revenge and hatred' (GM I:13). The particular intensity/frequency/duration of certain emotions serve to create cognitive dissonance for the slaves, and threaten to undermine the slaves' person-facing beliefs, in particular.

I will then suggest that in the *Genealogy*, Nietzsche gestures towards one possible means of resolving this tension between affective states and person-facing beliefs, namely via the mechanism of pretense. Pretense here is broadly understood as an imagined-state, which plays an intermediary role between the subject's affective states and their beliefs. These pretense-states allow for a second-order reinterpretation of emotions like vengefulness, such that they are re-framed and brought into line with the subject's self-deceptive beliefs. In the *Genealogy,* these pretense-states take the form of rich, vivid narratives (GM I:15), and provide a context in which their emotions are deemed acceptable and in keeping with their self-deceptive person-facing beliefs.

In this way, pretense-states tie in with Nietzsche's remarks on the importance of falsehoods and 'saving illusions' (BT; GS 107; GM III:24). As Anderson (2005:186) suggests 'Nietzsche's direct concern is not the existence or possibility of truth and knowledge, but their *value*'– in this way, imagined-states may displace knowledge or true belief as more appropriate or valuable. In the *Genealogy* pretense-states enable individuals to endure in the face of adversity, and play a therapeutic role, allowing for self-regulation in the subject's mental economy.

4.1 Self-Deceptive Beliefs in the *Genealogy*

The first question we must ask is what kind of self-deceptive beliefs do we find in the *Genealogy*? One feature of the slave-revolt is that the slaves desire to feel morally superior to the nobles. Poellner (2004) and Elgat (2015) both focus on the idea that in order for the slaves to feel morally superior, they must genuinely believe that their new slavish values of humility, compassion, patience, moderation etc. are good in themselves, and not merely instruments for revenge (cf. Elgat 2015: 526). For these new values to take root, there must not be any sense that the slaves have knowingly tricked themselves into holding these values: 'one cannot feel superior to another on the basis of values one does not really care about oneself'[66] (Elgat:2015:526). Acknowledging that these new values are held as part of a plan for revenge, might result in the slaves feeling superior in the sense that they are intelligent enough to successfully trick the nobles. But this would not be a case of feeling morally superior – intelligence would not suffice in this scenario. To feel morally superior the slaves would need to believe that their set of values is morally good.

The slaves must therefore believe that they are motivated to hold a certain value because, and only because, that value is intrinsically good, and consequently ought to be held. But it is not just a matter of being self-deceived about one's motivations; the slaves must also come to believe that they are exhibiting, (or at least intending to exhibit), these values, i.e. they must really think that they are, (or aiming to be), humble, compassionate, patient, mild. In this way, it seems that we end up with two kinds of beliefs: (i) beliefs which are orientated around the values themselves – value-facing beliefs, and (ii) beliefs which are orientated around the agents who exhibit these values – person-facing beliefs.

In the *Genealogy*, it seems that both types of self-deceptive beliefs are important for enabling this feeling of moral superiority. These new values must be seen to be exercised

[66] It is worth emphasising that the superiority referred to here is *moral* superiority in particular. It is plausible that someone could feel superior to another for reasons which they would, when being honest, admit are arbitrary and fatuous. One such case might be the British class system. My thanks to Eliot Michaelson for raising this point.

practically so that the slaves feel they are fully invested in these values, not just that they are lauding them in a detached, abstract manner, and for this we need person-facing beliefs. It is worth noting that value-facing beliefs and person-facing beliefs do not always overlap, for example a slave could believe that patience is intrinsically good but also believe that they themselves are not patient. In such a case, the slave would consider himself to be falling short of this value or sinning. Thus, the person-facing belief would be interpreted in such a way that it would not be seen to undermine the value-facing belief that patience is intrinsically good.

Are there ever instances where either value-facing and/or person-facing self-deceptive beliefs might be called into question? Self-deception is notoriously fragile (Gendler 2010, Audi 1988); should a self-deceptive belief be doubted by the subject, there is an obvious risk that this self-deception becomes unstable. One way this might occur is by becoming aware of the fact that you hold contradictory beliefs. If we take value-facing beliefs present in the *Genealogy* as an example, the conflict between these beliefs could be as follows:

[A]

(A1) the slaves' beliefs about their newly adopted slavish values (e.g. humility) i.e. *'I believe that humility is intrinsically good'*

(A2) the slaves' beliefs about their motivation for holding these values, i.e. '*the reason I believe humility is good is because I can use this value to exact revenge on the nobles*'

The idea in [A] is that an awareness of (A2) would threaten (A1) in that the realisation that the reason one holds value x is primarily instrumental would undermine the belief that value x is intrinsically good (and should be held only for that reason). So, if the slave comes to see that the only reason he is propagating and committing to a virtue like humility is because this activity provides him with a means of gaining some sort of superiority over his oppressors, thus functioning as a form of revenge, then it becomes apparent that he does not view humility as being purely of intrinsic worth. Consequently,

the belief in (A1) would cease to have the force it previously held in the slave's psychology.

Is there any way this kind of conflict could be avoided? One way [A] would not see any conflict in the subject's psychology is by understanding it as a case of (motivated rather than accidental) mistaken belief. Deflationary accounts[67] pursue this line of thought and characterise self-deception as systematic error due to a biasing factor(s), which operates unbeknownst to the subject.

In the case of the slaves, they erroneously believe (A1) even though the state of affairs in (A2) is actually the case - they are simply unaware of their true motivations. It seems that in this scenario the slave could quite happily maintain their mistaken belief, and (A2) drops out of the picture. This lack of self-awareness seems plausible – nowhere in the *Genealogy* do we find evidence that the slaves possess the kind of self-referential belief we find in (A2) – in this way (A2) appears far too sophisticated for the slaves. Instead we can replace (A2) with a relevant biasing factor (*), namely a desire to feel morally superior:

[A*]

(A1) the slaves' beliefs about their newly adopted slavish values (e.g. humility) i.e. *'I believe that humility is intrinsically good'*

(*) the slaves' desire to feel morally superior to the nobles

Elgat (2015) refers to Mele's deflationary account (2000) which explains self-deception via certain motivational biases, i.e. a desire regarding p or related to p, which causes the

[67] Deflationary accounts do not require any intention of the part of the subject to engage in self-deception, nor do they require the subject to be in two contradictory mental states. This is in contrast to intentionalist accounts of self-deception which broadly take self-deception to be modelled on interpersonal deception whereby the intention to bring it about that the subject believes p must be present (despite ¬p being the case).

subject to view and evaluate evidence in a biased way, resulting in a false belief. The resulting false belief is genuinely held by the subject. In Mele's account such desires will trigger two specific biases: (i) a confirmation bias and, (ii) a vividness of information bias. With the former, when the subject S possesses a desire related to p, S will tend to look for, both in memories and real life, confirming instances rather than disconfirming instances for p, and these confirming instances will be recognised more readily by S. With the latter kind of bias, S is more likely to recognise, attend to, and recall vivid data than pallid data, meaning that vivid data related to p will have a disproportionate influence on S's formation and retention of beliefs. Either or both of these biases will, according to Mele, be enough to structure and sustain a false belief.

How would this apply to the case of self-deception in the *Genealogy*? Elgat suggests that in the case of the slaves, a desire, namely to feel morally superior to the nobles, triggers these motivational biases, and we may speculate that:

> under the influence of the confirmation bias the slaves would tend to be on the lookout for evidence that support their desired belief and would thus gladly seize upon and be satisfied with the priest's teachings while downplaying or ignoring any countervailing evidence (their objective miserable predicament, the fact that throughout their lives they adhered to masterly values, etc.). Furthermore, the slave's desire for superiority would cause the priest's teaching to appear especially vivid in a way which would powerfully influence the slave's formation of beliefs.
>
> (2015:536)

In this way, the slaves are able to misinterpret evidence, such that the facts seem to support their desire to feel morally superior. This evidence, as Elgat notes, comes in the form of the priest's teachings. Alongside the slaves, Nietzsche introduces the character of the priest, who through religious teachings, initiates the reinterpretation of the status quo:

> The wretched alone are good; the poor, impotent, lowly alone are good; the suffering, deprived, sick, ugly alone are pious, alone are blessed by God, blessedness is for them alone – and you, the powerful and noble, are on the

> contrary the evil, the cruel, the lustful, the insatiable, the godless to all eternity; and you shall be in all eternity the unblessed, accursed, and damned!
>
> (GM I:7)

The priest's rhetoric of the oppressed as good and the oppressors as evil, thus offers the slaves a confirming instance of their belief that they are superior to the nobles: 'the lambs say to each other, 'These birds of prey are evil; and whoever is least like a bird of prey and most like its opposite, a lamb – is good, isn't he?'' (GM I:13). These teachings are then accepted due to the slaves' biases. In this way (A2) is replaced by (A*), i.e. the desire to feel superior to the nobles, and we avoid the conflict between (A1) and (A2). The slaves therefore come to genuinely believe (A1) without realising that they lack good reasons for doing so.

The deflationary account outlined in [A*] seems able to avoid a characterisation of the slaves as holding conflicting belief-states, and thus avoids one instance of cognitive dissonance[68]. But there is a more pressing concern. It is not clear that this deflationary approach acknowledges the full force of Nietzsche's descriptions of the slaves as being intentionally deceived. Nietzsche outright refers to the phenomenon as a 'sublime self-deception' [*sublime Selbstbetrügerei*] (GM I:13) in which the subject is far from being an innocent party: 'the man of *ressentiment* is neither upright nor naïve nor honest and straightforward with himself. His soul squints' (GM I:10). As Poellner (2004) notes, Nietzsche is insistent on this element of mendacity:

> The subject of *ressentiment* does not simply find herself, by the operation of some to her inscrutable mental mechanism, with 'moral values' which happen to be resources answering to her desires. Rather, the re-interpretation of her hatred of the object of *ressentiment* as a righteous disapproval of wickedness

[68] Another advantage of a deflationary approach is that it is able to avoid the static and dynamic paradoxes which often beset intentionalist models of self-deception. The static paradox focuses on the problem of simultaneously believing p and ¬p, whilst the dynamic paradox focuses on the apparent impossibility of the process in which a subject tries to hold a belief that p when they know or believe that ¬p. Appeals to split-mind models, sub-personal agency, or unconscious beliefs are often made by intentionalists in an attempt to avoid these paradoxes (e.g. Pears 1984, 1991, Davidson 1982, 1985a, 1985b).

> is the result of intentional action: hatred is 'masked' and 'counterfeited' (GM III, 14) as virtue…there is copious textual evidence that Nietzsche regards the subject of *ressentiment* as genuinely self-deceived.
>
> (2004:47)

One standard criticism which has been levelled against this sort of deflationary account of self-deception is that it becomes difficult to distinguish self-deception from wishful thinking (Porcher 2012). On this account to be self-deceived the subject only needs to hold a false belief that p, possess evidence that ¬p, and have some desire or emotion that explains why p is believed and retained. It is not clear this this fully captures the element of mendacity which Nietzsche emphasises. It is beyond the scope of this chapter to enter into this particular debate regarding ways of distinguishing deflationary accounts of self-deception from other forms of motivated believing such as wishful thinking, but I suggest that this worry is particularly pertinent for the *Genealogy* since Nietzsche makes clear that this is an instance of self-deception, rather than wishful thinking.

So, what are the slaves attempting to intentionally 'conceal' or 'mask'? As previously noted, it does not appear that the slaves possess the degree of self-awareness required for a belief such as that in (A2), i.e. a belief about their motivation for holding these values, i.e. '*the reason I believe humility is good is because I can use this value to exact revenge on the nobles.*' But one thing which they are aware of is their phenomenology, i.e. the emotions they experience.

As Poellner (2004) notes, the potential tension(s) in the slave's psychology could include not only belief-on-belief tension, but also tension arising due to affective mental states. Poellner offers a Sartrean-inspired account to accommodate a notion of intentional deception vis-à-vis belief-states, which in part draws on a distinction between unacknowledged, implicit conscious contents and implicit attitudinal character of experiences on the one hand, and conceptually articulated character and contents on the

other.[69]In an example like [A] the belief-states (A1) and (A2) are retained but in such a manner that no cognitive dissonance arises for the subject, and the self-deception remains intact.

However, as Poellner himself notes 'while the Sartrean approach has a good answer to the question how *ressentiment* can consistently involve a project of concealment (6), it is not clear how it can plausibly account for what is concealed' (2004:59). In the self-deceiving *ressentiment*-driven subject, what is concealed are not just beliefs, motivations, or intentions, but crucially emotions.

Since it seems there is no evidence to suggest that the slaves possess self-referential beliefs regarding their motivations, the slaves' emotions are therefore the most pressing problem vis-à-vis the stability of their self-deception. These emotions are of such a palpability that they cannot go unnoticed by the subject, (as would be required by the Sartrean approach). Poellner flags this concern:

> What could the method of concealment or 'masking' (GM III:14) of a present conscious state with high affective intensity possibly consist in, given that such states are necessarily in the foreground of consciousness, conspicuously manifesting their presence?
>
> (2004:60)

Elgat also briefly raises this worry: 'it would be implausible to hold either that they are unaware of the raging of *ressentiment* inside them or of how such strong emotions could cloud one's judgement and influence one's thinking' (2015:538). It seems that if we turn

[69] On this picture, something like (A2) would not be noticed or acknowledged by the subject, but is in principle noticeable via a shift of attention: 'those contents of consciousness we cannot acknowledge without realizing our practical inconsistency are indeed conscious, but pre-reflectively: they are unattended yet nevertheless skilfully negotiated 'background'. Avoiding explicit awareness of them is no more unproblematic than avoiding bumping into a piece of furniture without paying attention to it while perambulating engaged in intense conversation in a familiar environment' (2004:58). As to the question of what prevents the subject from shifting attention to the implicit content, (and thereby becoming aware of it), a Sartrean-inspired approach involves the subject 'intentionally adopting a non-reflective stance vis-à-vis her own intentional behaviour in certain contexts; and this stance includes *itself* among the items it pre-reflectively resists being made the objects of critical reflection and detection' (2004:59).

our attention to some of the emotions present in the subject's psychology, there is a very real chance of cognitive dissonance.

This section has examined the possibility of belief-on-belief tension arising in the slave's psychology, as outlined in [A], along with two different treatments of [A], one deflationary and one intentionalist. With both approaches, it seems cognitive dissonance can be avoided. However, as both Poellner and Elgat note, the real threat to the stability of the slaves' self-deceptive beliefs comes from their emotions.

In the next section, I will further investigate this worry and explore precisely how such cognitive dissonance can arise in the *Genealogy*, which emotions are involved, and the nature of their manifestation, i.e. whether their degree, frequency, and their duration makes a difference in generating this conflict in the subject's psychology.

4.2 Emotions in the *Genealogy*

How might emotions contribute to generating this cognitive dissonance, and threaten an individual's self-deceptive beliefs? We may first note that emotions need not necessarily create cognitive dissonance in the self-deceived subject: there are cases in which emotions serve to underpin and bolster a self-deception, and work to avoid any conflict in the subject's mental economy. Mele (2003) discusses the role emotions may have in driving a self-deception. He gives the example of the self-deceived husband, Bob, who believes his wife Ann is not having an affair. Bob experiences certain emotions which help to generate, and support what Mele understands to be his self-deceptive belief: 'Bob's love for Ann, or his fear that he cannot get along without her, may be a partial cause of his desire that she is not having an affair and, thereby, of his being self-deceived about this' (2003: 169). In such a scenario, it seems perfectly plausible that Bob's emotions would set to work generating confirmation and vividness of information biases such that Bob would misinterpret evidence, e.g. Ann's guilt regarding her affair manifests as her cooking an elaborate meal for Bob: Bob takes this gesture to be evidence of her love for him. Bob's emotions then are entirely consistent with his belief that 'Ann isn't having an affair'.

If we turn to the emotions described in the *Genealogy* however, it seems we do not have an analogous case. I suggest that the affective experiences Nietzsche describes are at odds with the person-facing beliefs held by the slaves– in these cases the emotions seem to threaten, rather than support the self-deception. Whilst Mele's deflationary account is attractive in that it can give an explicit role to emotions in generating, and sustaining self-deceptions, it is not clear how he would explain emotions which undermine the viability and stability of a self-deception.

So how are these emotions experienced, and how and when do they become problematic? Throughout the *Genealogy* Nietzsche is careful to emphasise the affective experiences of the slaves. As a result of their ambient environment, i.e. a constant state of oppression by the nobles, the slaves are described as feeling 'powerless' (GM I:10), 'miserable' (GM I:14), full of 'secretly smouldering emotions of revenge and hatred' (GM I:13) and 'rankled with poisonous and hostile feelings.' They experience 'a deep sadness' and cultivate 'a soil of self-contempt' (GM III:14). Any happiness the slaves experience 'manifests itself as essentially a narcotic, an anaesthetic, rest, peace, 'sabbath', relaxation of the mind and stretching of the limbs' (GM I:10). This suggests that their emotions associated with being oppressed are (i) the status quo (rather than acute, isolated episodes), and (ii) experienced in a palpable manner – the idea being that one is only in need of an anaesthetic if one is consciously experiencing pain. As Poellner notes, the descriptions Nietzsche offers of these emotional states: 'suggest strongly that the original negative affective state the subject 'sees' but 'want[s] not to see' continues to be 'visible', i.e. phenomenally conscious, to the self-deceiver (2004:48). We can therefore assume that at least some of these feelings of powerlessness, hatred, vengefulness etc. are experienced by the slaves in such a way or to such a degree that they are aware of them.

Whilst there is an extensive literature on the philosophy of emotions for example whether emotions are best understood via a cognitivist account (Lazarus 1991, Nussbaum 2004) or a perceptual account (Prinz 2006, Damasio 1994), for the purposes of this chapter, I will remain neutral on these debates, and focus on the particular viscerality of the emotions described in the *Genealogy*, suggesting that from Nietzsche's descriptions, the phenomenology of these emotions is of such an intensity that some kind of content is carried in them. I will examine which particular emotions threaten self-deception in the

Genealogy, and how the degree/frequency/duration of these affective experiences may undermine self-deceptive beliefs. The particular case studies I present are by no means exhaustive, (it is highly plausible that we may think of many different combinations of beliefs and emotions for the slaves), but are designed to show scenarios where conflict due to these affective experiences may not arise, and cases where conflict seems unavoidable.

Given the palpability of the slaves' affective experiences, how exactly can they generate cognitive dissonance and threaten their self-deceptive beliefs? In the previous section we drew a distinction between the slaves' beliefs which were value-facing and their beliefs which were person-facing.

If we first focus on a value-facing belief and the emotions of anger and humiliation experienced by the slaves, we have model [B]:

(B1) The slaves' beliefs about their motivations for holding their newly adopted slavish values, i.e*. 'I believe that humility is intrinsically good*'

(B2) The slaves' feelings of anger and humiliation

Here it seems that there does not necessarily need to be any conflict. It could be the case that the slave holds the belief in (B1) yet palpably experiences feelings of anger and humiliation. Particularly if, for example, anger is experienced as directed towards the nobles, this can be understood as anger at the fact that the nobles do not respect values like humility. In this scenario, this anger is seen as righteous anger, and as such there is no conflict between (B1) and (B2).

What if the anger is experienced in a more inchoate manner, i.e. the anger does not seem to be clearly directed at any particular target? Again, this need not conflict with (B1). Given the context of the priest's teachings, this kind of emotion could be interpreted as the emotion of someone falling short of their values, i.e. the sinner. In such a framework, qua sinner, the slave comes to understand these emotions as those associated with failure

and shortcoming, a lack of self-worth, guilt and shame, and so on, all the while still holding to his belief that a value like humility is intrinsically good.

In such a scenario (B1) and (B2) are not incompatible – the particular emotions experienced serve to reinforce the self-deceptive belief. The reason the 'sinner' feels so bad is because he holds certain values in such high esteem and aspires towards them.

The case of [B] thus poses no problem for Mele's account, and is in line with the role he accords to emotions in being able to drive and sustain a self-deception. The emotions in (B2) play a part in leading the slaves to select and/or misevaluate evidence to support their biasing desire (i.e. to feel superior to the nobles). As previously suggested, this evidence comes in the form of the priest's teachings. Given that the rhetoric of the priest's teachings plays upon feelings of humiliation - 'they tell me that their misery means they are God's chosen and select, after all, people beat the dogs they love best; perhaps this misery is just a preparation, a test, a training' (GM I:14), anger, and envy – 'one day *their* kingdom will come too' (GM I:15), it makes sense that the emotions in (B2) only serve to reinforce the slave's self-deceptive value-facing beliefs.

In this way, the emotions of anger and humiliation present in [B] do not create any cognitive dissonance for the subject and the self-deception is not threatened. We should note, however, that this works better with intermittent and relatively moderate experiences of anger – if anger is experienced constantly as high intensity rage then it is not clear that the sinner story could accommodate this. If the subject comes to realise that they feel this intense rage very frequently and for prolonged periods, they may well question what has prompted this and/or what it might indicate (plausibly something aside from falling short of virtues).

With [B] we have seen an example of how the slaves' emotions need not threaten their self-deceptive beliefs, in this case, value-facing beliefs. I will now examine a case [C] where the slaves' emotions do cause cognitive dissonance, and potentially undermine their self-deceptive beliefs. It seems that, in this scenario, the cognitive dissonance arises when we turn to the slaves' person-facing beliefs (which relate to an evaluation of their character and whether/how they manifest virtues), and focus on a different set of

emotions experienced by the slave, namely hatred, envy, and vengefulness. Here, the conflict could be as follows:

[C]

(C1) A belief the slaves hold about their character, i.e. *'I am humble'*

(C2) The slaves' feelings of hatred, envy, vengefulness vis-a-vis the nobles

This time it seems that a tension does arise between the beliefs in (C1) and the emotions experienced in (C2). Whilst on [B] the slaves' emotions did not pose a problem for their self-deceptive beliefs, if we shift firstly from focusing on value-facing beliefs like 'I believe that humility is intrinsically good' to focusing on person-facing beliefs like 'I am humble', and secondly from emotions of anger and humiliation to emotions of hatred, envy, and vengefulness, the self-deception becomes problematic.

In [C] these emotions seem to directly contradict the beliefs held in (C1). How can we square immediate and intense feelings of hatred and envy with a belief that one is meek and mild? Or acute feelings of vengefulness with a belief that one is kind and compassionate? It seems a huge stretch to try to understand these particular emotions within the framework of being a sinner, as we did in [B]. Even without any insight into their real motivations for holding these values, the very phenomenology of the slaves makes it difficult for them to avoid calling their self-deceptive beliefs in (C1) into question. It seems that self-deceptive person-facing beliefs are particularly at risk of being directly undermined by the subject's emotions.

How can a self-deception be sustained in the face of such high-intensity affective experiences which cannot be ignored by the subject? One way we might dissolve the tension in [C] is to draw a distinction between a character trait attribution like 'I am humble', and an episodic emotion attribution, i.e. 'I hate x on this particular occasion' or 'I feel envious of x on this particular occasion.' One common-sense understanding of character traits is that they are deep facts about a person which are relatively stable and enduring. Such traits are often not incompatible with episodes where individuals exhibit

behaviour which might suggest the contrary of that trait, e.g. a normally calm individual becoming annoyed when another driver cuts in front of them. Here, the behaviour of that individual would be seen as justified in an exceptional situation, and would not cause us to doubt the original assigned character trait.

In the case of the slaves, they might very well take a person-facing belief like 'I am humble' to be a character trait attribution, whilst also experiencing an emotion like hatred or envy episodically. This emotion would merely be viewed as a reasonable, yet uncharacteristic response to an extreme situation, i.e. being oppressed by the nobles who are so much better off than the slaves materially, socially, and politically. These responses could easily be understood as not conflicting with how the slave thinks of his character in general.

However, we might question whether certain emotions are easier to explain away in this manner than others. Whilst something like hatred or envy could be viewed as an uncharacteristic, and basic response in an extreme or exceptional situation, an emotion like vengefulness is potentially more problematic.

We might want to say that something like vengefulness belongs to class of emotions which are more sophisticated. It seems like vengefulness needs to be associated with a particular end, and is tied to very specific future-directed desires and intentions. These may preoccupy the subject to the level of obsession. Vengefulness also involves an element of planning, strategising, or plotting: the individual ends up partaking in some kind of project (whether in a real sense or an imaginary sense, e.g. a schoolchild being punished might daydream about how to get revenge on their teacher). In this way, the individual is forced to engage with this emotion in a way that need not be the case with something like anger. Hatred or envy may be experienced at the foreground of consciousness in a palpable manner, but need not necessitate that the individual engage in any kind of project – with vengefulness, structurally this emotion seems to call for such activity.

The option to explain away an emotion like vengefulness is still open but would require a far more complicated story as to why the individual can still retain character traits. For example if someone believes themselves to be a generally benevolent and forgiving person but repeatedly find themselves desiring and plotting revenge: here the

connotations of vindictiveness, spite, and insidiousness associated with vengefulness seem particularly problematic. So even if we attempt to dissolve the tension we saw in [C] by understanding (C1) as an enduring character trait and (C2) as an episodic emotion which is merely an uncharacteristic response to exceptional circumstances, this still necessitates some explanation or story to re-frame the emotion such that it does not undermine the trait observed in (C1).

The shift from [B] to [C] therefore demonstrates how the particular combination of beliefs and emotions affects whether the self-deception becomes threatened. It seems that with person-facing beliefs like 'I am humble' and the emotion of vengefulness (being a more sophisticated class of emotion) that we see a problem. Vengefulness does not look like it can be easily explained away with something like the sinner story in [B] or the character trait/episodic emotion distinction in [C]. For any tension to arise in the subject's psychology we would also have to assume that this emotion of vengefulness is experienced in an extreme manner, i.e. the experiences are of a high intensity, occur frequently, and last for significant periods of time. (Infrequent, mild, short-lived bouts of the emotion may go relatively unnoticed by the subject thus avoiding any cognitive dissonance). It appears that vengefulness, in particular, will need to be dealt with in some way to preserve the subject's self-deceptive person-facing beliefs.

Both Poellner and Elgat agree that high-intensity affective experiences must be managed in some way if the self-deception is to be sustained by the subject, and suggest that the solution may lie in reinterpreting or rebranding the affective content. How should we understand this notion of re-interpreting?

Elgat suggests that the entire project of the slave revolt in morality is itself revalued when the revenge and hatred present in *ressentiment* is reinterpreted as a call and hope for justice: 'as a consequence, the slave revolt in morality is now itself deemed as 'just': it is not an act of vengeance but is rather motivated by justice' (2015:539). In this way, he suggests that Nietzsche envisages the revaluation of *ressentiment* and the emotions involved as 'a meta-revaluation of the entire act of revaluation. It is thus a creative act of a higher order'. He then goes on to claim that with regards to the affective experiences of the *ressentiment* subjects, this process 'serves as a powerful (though not absolutely impenetrable) bulwark against… the threat that it poses to the viability of the slaves' self-

deception' (2015:540). Poellner also picks up on this potential solution of a re-branding of these emotions and suggests that we see evidence for such a phenomenon in (GM 1:14): 'what they hate is not their enemy, no! they hate 'injustice', they hate 'godlessness'.

How viable is re-interpreting emotions as a solution? Poellner raises a concern:

> This re-interpretation of negative affect as disapproval of 'immorality' makes possible its acknowledgement as well as its modification (into an experience of moral superiority). The problem of how current, conscious, emotional episodes like hatred could possibly be actively concealed from the subject therefore vanishes — they could not and do not need to be intentionally concealed. Yet, does this solution not merely replace one problem with another?
>
> (2004: 63)

Poellner goes on to argue that there are two problems with this possible solution of re-interpreting. Firstly, given that a reinterpretation of the emotions would require the subject to utilise concepts like humility, there is a circularity worry. He suggests that the *ressentiment* subjects do not have any real acquaintance with any of these new values such as humility, (presumably because this was not part of the prevailing value-set of the nobles), and consequently have no concept of the values to utilise in the modification and re-interpretation of their affective experiences. As such, it would seem that the subjects are not able to manage their cognitive dissonance.

One way that Poellner (2004:63) suggests we might overcome this worry is to consider two possibilities: 1) a second-hand acquaintance with the concepts - 'it is conceivable, for instance, that ….the subject utilises values which she has no acquaintance with, if she is aware that these values are recognised by *others* in her culture', or 2) the subject does have, in other contexts, a genuine (non-self-deceived), albeit relatively weak grasp of the values instrumentalised in *ressentiment.*

Poellner maintains that Nietzsche does not consider 2). He then goes on to argue that this problem undermines Nietzsche's claim in GM I:10 that '*ressentiment* itself turns creative and gives birth to values'. Instead we can only view *ressentiment* as a derivative or parasitic phenomenon 'that is dependent upon a prior recognition by someone of the

value it instrumentalises'. In order for the *ressentiment* subjects to utilise concepts like humility to reinterpret their emotional experiences, they must already recognise such concepts, having acquired them from elsewhere, rather than creating them themselves. So, on Poellner's account, not only are the slaves unable to reframe their emotions because they lack the concepts to do so, but we will have to rethink a key characteristic of *ressentiment* itself, namely that it is creative and can give rise to values.

However, in the next section, I will suggest that this dead-end may be avoided and that we may supplement both Poellner's and Elgat's account, by developing the idea of subjects reinterpreting their affective experiences such that the self-deception is preserved.

I suggest the following: (i) that the possibility Poellner offers in 2), i.e. that *ressentiment* subjects do have a genuine albeit weak grasp of these concepts is in fact accommodated for by Nietzsche, (ii) this is achieved specifically through the use of pretense-states, which through their narrative component enable the subject to entertain a relatively weak (and non-instantiated) but genuine grasp of the values, (iii) this grasp of the values allows the subject to modulate their affective states in such a way that they are brought into line with their self-deceptive beliefs.

The effect of these intermediary pretense-states therefore softens any tension between the subject's self-deceptive beliefs and their affective experiences, thus allowing the self-deception to be maintained. The next section will examine the nature of these pretense-states in the *Genealogy,* and how it is they provide subjects with a weak grasp of concepts such as humility.

4.3 Re-framing through Pretense

As outlined in the previous section, it seems that in the *Genealogy*, whilst *ressentiment* subjects may be able to avoid the kind of belief-on-belief tension we see in [A], the picture becomes more complicated once we factor in certain emotions. Specifically, it seems that a tension could arise between the subject's self-deceptive person-facing beliefs such as 'I am humble,' and some of their affective experiences such as intense, frequent, and

prolonged feelings of vengefulness. The conflict in [C] is such that it could undermine and threaten the stability of the subject's self-deception.

Due to the particular palpability and nature of an emotion like vengefulness (in that it forces whomever is experiencing it to engage with it in a way we do not encounter with an emotion such as anger), the subject's phenomenology alone may call their self-deceptive beliefs into question as it would be seen to directly undermine their beliefs about their moral character, i.e. why am I so overwhelmed by feelings of vengefulness if I am so humble /compassionate /patient /forgiving etc? This kind of tension would need to be alleviated in some way in order to allow the self-deception to be preserved.

In this section, I will suggest that one means of softening this kind of tension can be found in the *Genealogy*, namely, the use of pretense. I will begin by outlining what constitutes a pretense-state in the *Genealogy*, and how they are to be differentiated from belief-states. Following Gendler (2007), I will suggest that these pretense-states are generated by the imagination, and engagement with such states crucially involves an element of make-believe. However, I will depart from Gendler's account in that I do not believe, in the case of the *Genealogy* at least, that such pretense-states, qua pseudo-beliefs occupy the role normally assigned to belief in the subject's mental economy and governance of their actions. Moreover, as Porcher notes (2014:304), Gendler makes a very strong claim in saying that 'self-deception just *is* pretense, that is, the product of self-deception is pretense (a state that is belief-like but that falls short of constituting full-blown belief)' and that this is very different from 'allowing a *role* for pretense in the explanation of the *process* of (some forms of) self-deception.' I believe my account here advances this latter view in allowing a role for pretense-states in generating and sustaining a self-deception.

I suggest that these pretense-states: 1) specifically play an intermediary role between the subject's affective states and beliefs, and 2) modify the subject's affective states such that they are brought into line with the subject's self-deceptive beliefs. Modification of these affective states, however, should not be understood as an actual transformation in kind of the emotion but rather as a second-order reinterpretation of it cf. Poellner:

> It is very important to bear in mind that the process Nietzsche is describing is not a transformation of an emotion of one type into a different emotion. It is not the case, for instance, that the hatred of an oppressive class is, because it

> cannot be expressed, actually transformed into a different emotion, say, a love of egalitarian justice' so not a transformation in kind.
>
> (2004: 48)

In this way, I suggest that the subject is able to reinterpret their phenomenology by reframing emotions via these pretense-states, and that the key feature of pretense-states which allows for them to play this role is their narrative and imaginative richness.

How should we understand pretense-states, and how exactly can we distinguish them from beliefs? For the purposes of this chapter, I will take pretense to be broadly construed as an imaginative episode, i.e. where the subject S forms a mental representation of p (where p can include both propositional and non-propositional content), regardless of whether p actually obtains in the real world. This mental state is to be distinguished from other mental states such as perception, supposition, remembering, believing, desiring and so on (although an imaginative episode can be prompted or driven by any of these states). S engages with this mental representation via the activity of make-believe (cf. Gendler 2007:235), and this imagining can lead to other mental states, in particular affective experiences. In the *Genealogy*, I will suggest that the relevant pretenses take the form of rich, vivid narratives.

Since our concern here is how the subject's self-deception can be retained in the face of emotions which directly call some of those self-deceptive beliefs into doubt, we must ask why this necessitates a pretense, rather than, say, another belief. I suggest in this section, that the key function of pretense-states in the *Genealogy* is to soften tension between a subject's self-deceptive beliefs and their affective experiences by modulating and reframing emotions. But could this role also be performed by a belief? In the case of the slaves a belief that 'I am righteous' could plausibly reframe feelings of anger as experiences of righteous anger. Such an emotion suddenly seems justified. In this way, we could say that another belief-state would be sufficient for reframing the emotions.

However, there are reasons for thinking that a pretense-state will be more effective in bolstering a self-deception than an additional belief-state. This is because of an important distinction between pretense and belief. A first pass at setting up such a distinction might be to suggest that pretense-states demonstrate a certain robustness to evidence, (so that

these countervailing emotions would not affect the subject imagining that p), whereas beliefs are truth-tracking, and will be revised or dropped in the face of any counter-evidence.

However, this does not always appear to be the case. For example, we might think of an alcoholic, who, when questioned about their excessive drinking, comes up with a series of beliefs to explain their drinking on that particular occasion, e.g. 'I found out my wife was cheating on me/ I had a tough day at work/ I had to wine and dine a client/ I had some trouble sleeping last night.' These beliefs would appear to be robust against evidence, i.e. the alcoholic neglects to factor in the repeated occasions where he had lacked any real 'reason' to drink, or the fact that his boss/friends/family/spouse have all expressed concern over his excessive alcohol consumption. In such a scenario, if the alcoholic is challenged on one particular belief, and encouraged to revise it, another belief can simply take up the slack, and he does not need to face up to his alcoholism. In this way, a series of beliefs may function in such a way that they are robust to any countervailing evidence. So feasibly, at least in some cases, beliefs can possess the same characteristic – robustness to evidence- that pretenses can.

However, in terms of the impact on our phenomenology, it could be the case that on certain occasions a belief is simply too thin. Pretense has been noted for its ability to generate affective responses.[70]Moran (1994) contrasts imagining with mere supposition and suggests that the former has a tendency to give rise to affective episodes where the latter does not, and the literature on fictional emotions documents how we are able to have strong emotional responses to imagined situations. In terms of differentiating between the affective response generated by a belief compared to those generated by a pretense, we might compare our emotional reaction to a belief that slavery is bad with our emotional reaction to watching a film about slavery which involves our imaginatively engaging with graphic and harrowing content. It would seem that in the latter case, our affective response would be heightened after having been forced to imaginatively engage

[70] We might compare this to Gendler's notion of alief (2008a; 2008b) where the subject exhibits the kind of emotional and behavioural responses typically associated with belief.

with the narrative. The 'simulative off-line processing' (Gendler:2007:235) that the characters and plotline provide enable us to emotionally engage in a way which might not be possible with the mere belief that slavery is wrong. The effect on our phenomenology in such a case is stronger.[71] The important difference between a pretense and belief for our purposes, therefore seems to be its imaginative, narrative richness.

So, do we find these sorts of pretenses in the *Genealogy*? Nietzsche seems to refer to the imaginative capacity of the slaves at several points throughout the text. Throughout the first essay Nietzsche emphasises this role of imagination vis-à-vis the slaves' response to their oppression. In GM I:10 it is made clear that the slaves are relatively impotent in terms of actions but are able to draw on their imaginative resources to come up with a response to their oppression: 'the *ressentiment* of those beings who, denied the proper response of action, compensate for it only with imaginary revenge', and 'the distortion with which the entrenched hatred and revenge of the powerless man attacks his opponent – in effigy of course'.

This 'effigy' is part of the imaginary world the slave constructs – Nietzsche uses the verb '*sich ausdenken*' which can be translated as 'to imagine', 'to dream up', 'contrive', 'invent', 'think up.' There is a clear emphasis on the role of imagination here: 'he has conceived of the 'evil enemy', '*the evil one*' as a basic idea to which he now thinks up a copy and counterpart', the 'good one' – himself!' (GM I:10). In this way, the slave effectively dreams up heroes and villains, and in doing so creates an identity for himself – the wronged hero.

Later in GM I:13 Nietzsche again uses *sich ausdenken*, suggesting that the slave comes up with the notions of good and evil which correspond to these heroes and villains: 'of good as thought up by the man of *ressentiment* [*sich ausgedacht hat*].' Here Nietzsche suggests that the slave constructs his own inner world complete with value systems.

[71] To return to the alcoholic example, it might be objected that a series of beliefs could generate the same affective impact. However, by this stage, if the alcoholic has come up with such an intricate, richly embellished network of beliefs which has no bearing on the real state of affairs, arguably this in itself becomes a pretense, as the alcoholic is forced to play through this network.

One term which reinforces the idea that the *ressentiment* subjects make use of their imaginative capacity is '*Vorwand*.' Often translated as 'pretext', '*Vorwand*' can also be rendered as 'pretense' or 'make-believe.' In GM III:15 Nietzsche claims that the *ressentiment* subject is 'frighteningly willing and inventive [*sind allesammt von einer entsetzlichen Bereitwilligkeit und Erfindsamkeit*] in their pretexts for painful emotions.' The mention of '*Erfindsamkeit*' here is pertinent – someone who is '*erfinderisch*' is creative, innovative, imaginative, and Nietzsche suggests that the slave meets these criteria- they mischaracterise present scenarios, imagining that they have been wronged or they mine their memories to look for 'stories' which will allow them to be suspicious, (or exercise other 'reactive' emotions), and view others as potential causes for their suffering. This is achieved through the use of narratives, none of which correspond to the actual states of affairs, but which are designed to alleviate his suffering:

> [T]he man of *ressentiment* is neither upright nor naïve, nor honest and straight with himself. His soul *squints*; his mind loves dark corners, secret paths and back-doors, everything secretive appeals to him as being *his* world, *his* security, *his* comfort.
>
> (GM I:10)

Here, Nietzsche seems to suggest that the man of *ressentiment* is somewhat reality-aversive, preferring to construct fantasies which will distract from the real state of affairs, and later in GM I:14 he even refers to them having a 'phantasmagoria' of 'anticipated future bliss' [*ihre Phantasmagorie der vorweggenommenen zukünftigen Seligkeit*].

Nietzsche traces the propensity of the *ressentiment* subjects to generate these pretenses to the fact that they are fundamentally unhealthy:

> These worm-eaten physiological casualties are all men of *ressentiment*, a whole, vibrating realm of subterranean revenge, inexhaustible and insatiable in its eruptions against the happy, and likewise in masquerades of revenge and pretexts for revenge: when will they actually achieve their ultimate, finest, most sublime triumph of revenge?
>
> (GM III:1)

This link is important because it supports the idea that at no point are the original emotions actually changed or removed altogether from the subject's psychology. If this were the case, it would suggest that the slaves would be far less sick and problematic than Nietzsche maintains: 'the *sickly* are the greatest danger to man: *not* the wicked, *not* the 'beasts of prey' (GM III:14). The very fact that the slaves utilise mechanisms such as self-deception and 'lie themselves into' things as opposed to being 'frank and confident' with themselves as the nobles are - [*während der vornehme Mensch vor sich selbst mit Vertrauen und Offenheit lebt*] - is for Nietzsche, symptomatic of their sickliness.

Despite their lies, the slaves remain 'poisoned' by their emotions. We can therefore see why, as Poellner makes clear (2004), no real transformation in kind occurs vis-a-vis these emotions, e.g. from hatred of the oppressors to love of equality:

> [I]t is quite clear that no transformation of this kind is supposed to occur in *ressentiment*. Rather, the original negative emotion towards a 'not-self' continues to motivate the subject, but it is 'masked', 'counterfeited', 'lied into' (*umgelogen*, GM I, 14) something it is not, namely, love of some putative good — humility, prudence, justice — for its own sake, and a consequent disapproval of moral 'evil'.
>
> (2004:48)

The emotions of anger, humiliation, hatred, envy, and crucially vengefulness therefore all remain in the *ressentiment* subject's mental economy.

How precisely do pretenses work to re-frame these emotions so that they are 'lied into' something they are not? I suggest that the process is twofold. The narratives constructed: 1) provide the slaves with a weak grasp of concepts such as humility, thus addressing the concern Poellner raised regarding whether the slaves have any real acquaintance with the concepts needed to reinterpet their affective experiences, (we might think of passages from the Bible such as Chronicles 1. 7-12, 7.11-15 where we get narratives extolling humility), and 2) provide the slaves with a context in which their reactive emotions are deemed acceptable (for example, Ephesians 4:26-27).

The imaginary simulation of revenge present in these narratives provide the slaves with something close to fictional emotions thus enabling them to actively engage with, rather

than repress or remove emotions of hatred, anger, vengefulness. The combination of these two factors work to ease any cognitive dissonance for the slaves – they are able to reinterpret their emotions as justified, i.e. these are the emotions experienced by 'the virtuous' and 'the blessed' who have been wronged.

In GM I:14 Nietzsche describes the slaves as self-styling themselves as 'the blessed,' who possess virtues of humility, compassion, patience and so on:

> [A]nd impotence which doesn't retaliate is being turned into "goodness"; timid baseness is being turned into "humility"; submission to people one hates is being turned into "obedience" (actually towards someone who, they say, orders this submission – they call him God). The inoffensiveness of the weakling, the very cowardice with which he is richly endowed, his standing-by-the-door, his inevitable position of having to wait, are all given good names such as "patience", also known as the virtue; not-being-able-to-take-revenge is called not-wanting-to-take-revenge, it might even be forgiveness.
>
> (GM I:14)

Effectively this role-playing provides the slaves with a (non-instantiated) and weak grasp of concepts such as humility. This character of the virtuous blessed individual is also echoed later in GM III:15 with mentions of 'martyrs' and a 'saint' who speaks 'as meekly as a lamb'.

Poellner's concern was that the slaves did not possess any real acquaintance with concepts like humility, and that without this they would not be able to reinterpret their affective experiences. However, we may object that the notion of acquaintance Poellner is relying on here is too strong – it is not obvious that a weaker acquaintance would not be adequate for the slaves to reinterpret their emotions. It seems that the particular imagined narratives the slaves generate provide enough acquaintance, (albeit very weakly), with concepts like humility or patience, such that they are able to utilise these concepts to understand their emotions as those exhibited by the wronged, virtuous individual. Just as imaginatively make-believing or daydreaming that you have a pet unicorn or that you can fly, provides you with a weak and non-instantiated concept of 'unicorn' or 'flying human', so might imaginatively pretending that you are humble or

patient provide you with a weak concept of these values even though they do not happen to be instantiated in real life.

The other key effect of these narratives is that they provide an acceptable context for the slaves' reactive emotions. Nietzsche provides us with a clear example of this in GM I:15 with a passage from the 'enraptured visionary' [*der entzückte Visionär*] Tertullian's *De spectaculis*. Nietzsche introduces the section by referencing Dante's *Inferno* and suggesting that a more apt inscription to stand over the gateway to 'Christian Paradise' would be 'Eternal hate created me as well.' Here, we get a sense of the slave's emotions playing an important role in structuring and creating their visions and ideals, as Nietzsche hints heavily that we ought to equate this 'paradise' with these emotions: 'For what is the bliss of this Paradise? . . . We might have guessed already.' Nietzsche then frames the Tertullian passage by placing it directly after a quotation from Aquinas' *Summa Theologica*: 'the blessed in the heavenly kingdom will see the torment of the damned so that they may even more thoroughly enjoy their blessedness' –suggesting that we are about to see an instance of how the slaves gain relief and satisfaction through these imaginary narratives, and more precisely, narratives of revenge.

The excerpt from Tertullian depicts a spectacle of torment of the damned with the faithful gleefully witnessing their persecutors' punishment:

> At which one shall I gaze in wonder? At which shall I laugh? At which rejoice? At which exult, when I see so many great kings who were proclaimed to have been taken up into heaven, groaning in the deepest darkness together with those who claimed to have witnessed their apotheosis and with Jove himself.[72]
>
> (GM II:15)

The descriptions become progressively more graphic: 'and when I see those [provincial] governors, persecutors of the Lord's name, melting in flames more savage than those with which they insolently raged against Christians!' and sadistic: 'then the tragic actors will

---

[72] '*Quid admirer! Quid rideam! Ubi gaudeam! Ubi exultem, spectans tot et tantos reges, qui in coelum recepti nuntiabantur, cum ipso Jove et ipsis suis testibus in imis tenebris congemescentes*'

be easier to hear because they will be in better voice [i.e. screaming even louder] in their own tragedy. Then the actors of pantomime will be easy to recognize, being much more nimble than usual because of the fire.' The direct reference to 'the imagining spirit' implies that it is through their imagination that the slaves will be able to exert power over their oppressors.

This imagined scenario thus provides a cathartic outlet for the slaves' feelings of vengefulness. They are able to execute strategies for revenge with great precision and detail, the possibilities for doing so are plentiful (Tertullian speaks of an 'ample breadth of sights' where each and every one of the persecutors is punished), and the timeframe – a 'final and everlasting day of judgement' ensures that this opportunity for revenge is never exhausted.

Vengefulness in particular is thus encouraged and facilitated in this narrative, alongside other reactive emotions of hatred, resentment, envy, and anger. Since the narrative is able to encompass the whole range of these reactive emotions, it emerges as a particularly effective means of reframing the slaves' affective experiences. It seems that with the repeated mention of 'imaginary revenge' throughout the first essay, which is then writ large in the Tertullian passage, Nietzsche is outlining something close to fictional emotions. Though the narratives or pretexts which the slaves construct are imaginary, they still serve to elicit emotional responses and in doing so, provide the slaves with their 'anaesthetic' for their suffering (GM I:10). These emotions are not repressed or removed, but instead expressed via these pretenses- in this way the narratives play a therapeutic role for the slaves.

The combination of allowing the slaves to style themselves as 'the virtuous' whilst expressing their reactive emotions, particularly vengefulness, works to ease any cognitive dissonance in their psychology. Person-facing beliefs that we saw in [C] such as 'I am humble' are able to co-exist alongside emotions of vengefulness, anger, hatred and so on, even when these are experienced as being of a high intensity, a high frequency, and a long duration by the subject. Pretense-states enable these self-deceptive beliefs to remain intact, and sincerely believed by the slaves. The modification of emotions like vengefulness therefore consists in these emotions being transferred to an acceptable

context where they are deemed to be justified: for this, the narrative element of a pretense is required.

It would be implausible to suggest that the *ressentiment* subjects are continuously engaged in these pretense-states (doing so would be tantamount to being delusional, and there is no textual evidence in the *Genealogy* for such a claim), but it seems that even occasional engagement in these pretenses could be sufficient to modify the subject's psychology, and it does not seem to be a huge stretch to suggest that in instances where the slaves might be experiencing cognitive dissonance, they might be more likely to indulge in these pretenses.

Conclusion

In conclusion, I have suggested that the self-deception we find in the *Genealogy* emerges as a case of self-regulation (in weak, sick individuals). With the examples of [A], [B], and [C] we have seen how the tension in the slaves' psychology comes not from a conflict between belief-states, but rather from the presence of certain emotions (vengefulness in particular) experienced with a specific frequency/intensity/duration.

I have suggested that the use of pretense-states provides the slaves with an acceptable outlet for their emotions such that they do not cause the kind of cognitive dissonance which may conflict with their self-deceptive beliefs. The pretense-states are able to achieve this thanks to their narrative richness, and as such, are more effective than, for example, additional belief-states in easing cognitive dissonance.

Avoiding such cognitive dissonance is crucial for allowing the slaves to maintain a feeling of moral superiority over the nobles, and accordingly, to endure and prosper (as far as their circumstances permit). Self-deception in this scenario thus acts as a form of therapy for the slaves. In this way, we see how psychological self-regulation can involve capacities such as imagination, and that the kind of emotional simulation this provides, allows the subject to reinterpret their phenomenology.

5. Self-Regulation and Nietzschean Health

Introduction

When sketching a theory of a Nietzschean self, and how it might operate, one way to approach this might be to consider how such a self responds to, and copes in challenging contexts, i.e. how it self-regulates. Broadly construed, we can understand self-regulation as the maintenance or re-calibration of some kind of balance within the self, physiologically or psychologically, particularly in the face of adverse circumstances. Considering the operations of a self under this lens of self-regulation enables us to view the self, not as a static entity, but rather as something that is responsive and dynamic. Nietzsche himself speaks of an 'instinct of self-restoration' (EH 'Why I am so Wise' 2) – this seems to imply that the self has some ability to sustain itself and recuperate.

How can we understand self-regulation on the Nietzschean picture? A first pass might be to consider whether it is reducible to other values such as strength or flourishing. Nietzsche frequently refers to, and praises these notions throughout his works (e.g. GM I:10), and it might seem that a self that can self-regulate is a self that is strong or a self that is flourishing. However, there are problems with this approach: by Nietzsche's lights, both strength and flourishing are value-loaded terms, and discussions of these terms have a distinctly elitist flavour. Nietzsche seems adamant that not everyone is strong, for example differentiating between the weak herd and the strong artistic types, and he often seems to conceive of flourishing in a particularly narrow way; aligning this activity with artistic capacity, aesthetic excellence, and culture [*Kultur*] ('The Greek State'). Self-regulation, by contrast, seems to be something that everyone engages in (albeit, perhaps, to greater or lesser extents, and with more or less success). One obvious example would be the slaves in the *Genealogy*: these individuals self-regulate in that, through their self-deceptions and pretense-states (see Chapter Four), they find some means of enduring their difficult circumstances. Although they self-regulate, Nietzsche by no means characterises the slaves as strong or as flourishing. For this reason alone, at least, it seems that self-regulation and notions of strength or flourishing could quite easily come apart.

The latter notions are presented as being commendable, and exceptional, whilst the former appears to be far more commonplace.

How else might we then characterise self-regulation? The idea that an individual recalibrates and recuperates through self-regulation may lead us to align this activity with health. An ability to recover and convalesce after challenging periods would seem to be an important facet of health. We tend to think of a healthy individual as someone who is not only resistant to illness in the first place, but who, should they succumb to illness, has a propensity to regain their health quickly and easily. Understanding self-regulation in terms of health may then seem like a natural option.

However, as I will argue, on the Nietzschean picture, self-regulation emerges as a necessary but not sufficient condition for his distinctive understanding of health, and that there are occasions where these notions come apart. The first section will review attempts made by commentators to understand Nietzsche's concept of health, and use them to draw out some of the differences between health and the activity of self-regulation. The second section will then use the example of the slaves in the *Genealogy* as a case-study of self-regulation to fine-tune this notion. Here we have a particularly complicated example in that these individuals seem to exhibit some of the marks of health observed in the first section, yet are classified by Nietzsche as fundamentally sick. For this reason, they present an especially interesting case of very efficient self-regulation, and demonstrate how this notion is a distinctive capacity in its own right.

Triangulating self-regulation with health in this way will therefore enable us to begin to understand self-regulation as an independent notion, whilst also providing us with a more fine-grained understanding of Nietzschean health. We will see that there are various aspects of self-regulation which emerge, such as endurance, and efficiency, which (perhaps surprisingly) are not sufficient for health.

5.1 Defining Health

In order to understand how self-regulation and health may come apart, we must first begin by trying to understand what Nietzsche takes health to be. This section will review two approaches to defining Nietzsche's concept of health. The first approach tries to understand health in terms of other Nietzschean notions (Letteri 1990; Neuhouser forthcoming), and the second approach makes reference to a) the arrangement of drives within the self, and b) the quality of these drives, i.e. whether an appropriate drive is in authority within the drive hierarchy (Huddleston 2017). I will then use these approaches to highlight some differences between health and self-regulation.

The first thing to note is that Nietzsche's understanding of health is complex: throughout his writings (but especially in his later works), Nietzsche appears to think in terms of 'healthy' or 'unhealthy', applying these terms to entities as varied as individuals (GM II:15, EH 'Why I am so wise' 1), entire cultures (HAH I:224; A57), drives (GM III), values (TI 'Morality as Anti-Nature' 4), and religions (A51; TI 'Improvers of Humanity' 3). The idea that 'healthy' is applied in all of these cases initially appears confusing - how are we to make sense of this term being applied to such diverse phenomena? One element of Nietzsche's thought to keep in the background is his so-called 'naturalism', and his claim in BGE 230 to 'translate man back into nature.' With this in mind, we can begin to see how Nietzsche might place the physiological and psychological, i.e. components of a self, in some kind of continuum with, for example, the moral or the aesthetic. In this way, Nietzsche is introducing an idiosyncratic notion of health.

In *The Gay Science,* we find Nietzsche's most protracted discussion of health (GS 120; GS 382) – this has been the main focus for many commentators (Letteri 1990). Here, Nietzsche addresses two questions: 1) whether it is even possible to define health and, 2) 'whether we can really dispense with illness?' It is worth examining the first passage (GS 120) in full:

> Health of the soul.- The popular medical formulation of morality that goes back to Ariston of Chios, "virtue is the health of the soul," would have to be changed to become useful, at least to read: "*your* virtue is the health of *your* soul." For there is no health as such, and all attempts to define a thing that way have been

> wretched failures. Even the determination of what is healthy for your body depends on your goal, your horizon, your energies, your impulses, your errors…and above all on the ideals and phantasms of your soul. Thus there are innumerable healths of the body; and the more we allow the unique and incomparable to raise its head again, and the more we abjure the dogma of the "equality of men", the more must the concept of a *normal* health, along with a normal diet and the normal course of an illness, be abandoned by medical men. Only then would the time have come to reflect on the health and illness of the *soul* and to find the peculiar virtue of each man in the health of his soul. In one person, of course, this health could look like its opposite in another person. Finally, the great question would still remain whether we can really dispense with illness-even for the sake of our virtue and whether our thirst for knowledge and self-knowledge in particular does not require the sick soul as much as the healthy, and whether, in brief, the will to health alone, is not a prejudice, cowardice, and perhaps a bit of very subtle barbarism and backwardness.

At a first pass, Nietzsche's declaration that 'there is no health as such, and all attempts to define a thing that way have been wretched failures' may lead us to view any such project as fundamentally misguided. On closer inspection however, as Huddleston (2017) has convincingly argued, to read Nietzsche as denying there is a concept of health *per se* would be mistaken- instead his claim is:

> that the valence of particular characteristics will be different in different cases. We need to interpret these characteristics in the context of one another in order to determine whether they redound to health or not. What may be a detriment to overall health in the case of one person may be a boon to overall health in the case of another. As such, we cannot infer from particular marks (or symptoms), regarded in isolation, that a person is healthy (or sick). Holistic interrelation of these characteristics is crucial.
>
> (2017:137)

In this way, the 'anti-universalist tenor' of this passage 'is not inconsistent with their being a core notion of health, understood in a suitably general way' (2017:137). This may vary across individuals and even within the same individual, i.e. what is 'healthy' when

one is a child may be different to what is 'healthy' when one is an adult cf. (KSA 9. 11 [112]): one simple example of this is the number of hours of sleep required according to age.

Read this way, it may then be tempting to collapse this notion of health into other concepts, and define various 'healths' as differing degrees of these concepts relative to the individual's contexts. What kind of concepts might these be?

The fact that Nietzsche immediately follows this claim by introducing the relation between health and illness, may lead us to focus on notions which would allow the individual to engage with this relation. These might be combative notions like 'overcoming' or 'struggle' where illness is battled and dispensed with. Or the notions might be more recuperative like 'reinterpretation', where illness is transformed or accommodated for in some way. Letteri (1990) opts for the former approach, and argues that 'for Nietzsche, health is essentially overcoming' (1990:410). Letteri understands the notion of overcoming against the background of Nietzsche's discussions of power, and argues that health as overcoming manifests against resistances or incapacities, i.e. illness. This interpretation, he maintains resists a 'view of health as a pristine state of inertia, as mere lack of disturbance' and instead proposes that 'health is…not simply the absence of weakness: it is, rather, the free admission of weakness into an arena of struggle with the aim of conquering it and thereby becoming stronger' (1990:411).

However, aligning health solely with overcoming is problematic for three reasons. Firstly, there is a worry that the notion of overcoming is perhaps too crude vis-à-vis Nietzsche's stance on sickness in GS 120. Given that Nietzsche often speaks of the value of suffering (BGE 225; GS 7; GS 13; GS 318; GS 325), it seems that he resists the idea that health should be defined as an absence of illness qua pain and suffering. (We should also note that not all forms of illness involve explicit experiences of pain, although they may involve degeneration or impairment e.g. a psychosis or dementia). There are also scenarios where pain serves as a useful warning sign to ensure we stay healthy: consider the case of a runner acknowledging that he should rest because of a twinge in his foot. Whilst he could keep running with reasonably minimal discomfort, this pain signals that to do so would cause complications further down the line. In this example pain actually helps us to preserve health. Health therefore seems to be something more than just the

definitional opposite of illness and vice versa. In *Ecce Homo*, Nietzsche speaks of the usefulness of illness, and how it has enabled him to refine his perspectives and become healthier:

> To be able to look out from the optic of sickness towards *healthier* concepts and values, and again the other way around, to look down from the fullness and self-assurance of the *rich* life into the secret work of the instinct of decadence - that was my longest training, my genuine experience.
>
> 'Why I am so wise' 2

As Janaway (2017:80) notes, suffering can be useful to the sufferer in that by:

> undergoing particular sufferings' they 'may re-configure their attitudes in ways that are important to them. They may understand things differently, begin to feel different emotions, be released from feeling other emotions, change their self-conception, develop new capacities and find a meaning in the course of events or in their life as a whole.

If we understand overcoming to be a case of eliminating illness (at least illness conceived of as pain and suffering) altogether, then this seems difficult to square with Nietzsche's remarks on the usefulness of illness. It seems that on Letteri's reading of overcoming, the idea is that illness is to be battled with before being eradicated – Nietzsche, however, seems to imply that there is a value to illness other than the fact that it is something to be wrestled with. In this way, he offers an understanding of health that is more nuanced than the notion of normal functionality undisrupted by illness.

Secondly, as Huddleston notes, we should resist placing too much weight on the characterisation of health as this idea of overcoming, and the agonistic relationship between health and illness should be restricted to what Nietzsche terms 'the great health' (GS 382):

> This form of health is closely bound up with sickness in that it is possible only through a certain *overcoming* of sickness. It is won through long, arduous struggle. As Nietzsche will put it, one does not just 'have' it, but 'acquires it continually' (GS 382). While it is important to bear this form of health in mind, I think it would be a mistake to extrapolate from this condition to something

> true of all spiritual health, let alone of all health. This close constitutive dependence between health and the overcoming of sickness (physical and spiritual) is a feature of a particularly exalted kind of health; indeed, the very fact that Nietzsche marks the form of health off in this way, with the special adjective 'great', would suggest that not *all* health is like this.[73]

Thirdly, I suggest that, as Nietzsche himself makes clear in *Ecce Homo*, the capacity to overcome illness is only possible if the individual is already fundamentally healthy:

> Something with a typically morbid nature cannot become healthy, much less make itself healthy; on the other hand, for something that is typically healthy, sickness can actually be an energetic *stimulus* to life, to being more alive.
>
> (EH 'Why I am so wise' 2)

We also find this thought in The Wagner Case: 'sickness can itself be a stimulus of life: it is just that you have to be healthy enough for this stimulus!' (WC: 5). To define health as essentially overcoming, as Letteri does, misses the mark somewhat. If, as Nietzsche states, you can only exercise this capacity to overcome illness, if you are already healthy, then it is not obvious that we should equate these two things. Health might be a necessary condition for the capacity to overcome illness, but this can imply that it is something over and above the capacity: overcoming is only part of the story. For example, in physiological terms, we might think of how someone with a strong immune system does not succumb to a winter flu – the virus enters their system, but because they are already healthy, their immune system is able to defeat or 'overcome' the virus. It seems we would want to say here that there is something about their immune system other than its ability to defeat the virus that designates it as healthy. Overcoming is certainly part of the story but there are other definable, more specific features we would want to refer to when explaining why that person is healthy.

---

[73] cf. 'Health and sickliness: be careful! The yardstick remains the body's efflorescence, the mind's elasticity, courage and cheerfulness - but also, of course, how much sickliness it can take upon itself and overcome - can make healthy. What would destroy more tender men is one of the stimulants of great health.' (Notebook from autumn 1885-1886, 2 [97]).

So, it seems that health as 'overcoming' or 'struggle' may not get us very far in defining Nietzschean health. Which other notions might be employed? One characteristic which Neuhouser (forthcoming) draws our attention to, is the notion of interpretation. When discussing Nietzsche's concept of biological health, Neuhouser beings by appealing to Nietzsche's ideas of power and life. (We should note that, as I will discuss in the final section of this chapter, it is not always obvious that Nietzsche's concept of health maps neatly on to life). Neuhouser suggests that life's aim is the process of continually striving to create conditions for discharging power: 'Nietzsche characterises life – animated by a will to life (*Lebenswille*)- as a series of activities that aims not primarily at self-maintenance or reproduction but at power – at creating ever "*greater* units of power", and maintains that illness is thus understood as impeding this particular function: 'illness involves some disturbance in, or blocking of, life's characteristic activity of creating ever more powerful configurations of life' (forthcoming:4). How then does this function of life operate? Neuhouser suggests that we can understand this function as interpretation, i.e. the assigning of meaning to seemingly 'random, unconnected, "meaningless" happenings':

> [T]o interpret, then – to give something a meaning – is to impose a function on what at first is merely 'there' by incorporating it into a system of purposeful activities such that it comes to serve an end of the organism as whole.
>
> (forthcoming:4)

and he draws upon the following quotation from the *Genealogy* to support this claim:

> every happening in the organic world is an *overpowering*, a *mastering*, and every overpowering….is itself a re-interpreting, a fitting into place, in which previous 'meaning' and 'purpose' must be obscured…or extinguished.
>
> (GM II:12)

Health thus consists in an 'order-imposing function', which can be reduced to this activity of interpreting. Neuhouser unpacks this idea at the level of the drives:

> [A] hierarchical organisation in which higher (or "nobler") functions rule over (or "dominate") lower functions (GM II:1), making the living being into a

> purposefully ordered whole – into, in other words, an organism, in which specialised functions work together to further the vital ends of the whole.
>
> (forthcoming:5)

The idea here seems to be that the drives will interpret events and experiences they encounter by imposing an order or meaning upon them, bringing them into the hierarchy, such that the individual is able to continue to discharge power. Neuhouser does not give a specific example of this but an obvious candidate to consider for this process would be the suffering which the *ressentiment* subjects encounter in the *Genealogy*. As a result of the political/social/material oppression the slaves and priests face, they experience a high degree of suffering which makes their situation seem irremediable and without hope. Nietzsche goes on to suggest that the brute fact of this suffering is not really the main problem: 'what actually arouses indignation over suffering is not the suffering itself, but the senselessness of suffering' (GM II:7) – it is the lack of meaning or purpose of the suffering that renders it intolerable:

> Man, the bravest animal and most prone to suffer, does not deny suffering as such: he wills it, he even seeks it out, provided he is shown a meaning for it, a purpose of suffering. The meaninglessness of suffering, not the suffering, was the curse that has so far blanketed mankind, – and the ascetic ideal offered man a meaning!
>
> (GM III:28)

Before the ascetic ideal, the suffering of the slaves and priests appears as one of those 'random, unconnected, "meaningless" happenings' which must be assigned some meaning or order for it to be rendered bearable. The ascetic ideal allows the *ressentiment* subjects to reinterpret their suffering as meaningful – it is now viewed as a consequence of sin (which can be cured), and is a necessary step to becoming 'good':

> Only those who suffer are good, only the poor, the powerless, the lowly are good; the suffering, the deprived, the sick, the ugly, are the only pious people, the only ones saved, salvation is for them alone, whereas you rich, the noble

> and powerful, you are eternally wicked, cruel, lustful, insatiate, godless, you will also be eternally wretched, cursed and damned!
>
> (GM I:7)

Here we have a sophisticated example of how meaning is assigned to experiences in order to generate order. This activity allows the *ressentiment* subjects to endure their suffering, and to make sense of its presence within their mental economy.

What is puzzling, however, is that Neuhouser characterises illness as including characteristics of 'powerlessness, passivity, reactivity, leadenness, and, perhaps most important, an incapacity to impose order, or meaning, on encountered realities' (forthcoming:7). Whilst we can certainly attribute the majority of these characteristics to the *ressentiment* subjects, it does not seem as though the last characteristic is a good fit. These individuals seem particularly adept at imposing order and meaning onto their encountered realities, so much so that an entire framework of a morality system is generated. So, going by Neuhouser's description of illness here, it looks like the *ressentiment* subjects do meet this condition of health, namely reinterpretation and order-imposing function.[74] Yet Nietzsche persistently refers to them as 'sick.'

One option might be to invoke the idea that there are degrees of health (and illness), and that the *ressentiment* subjects merely possess a very low degree of health – they still exhibit some capacity to interpret and impose order. However, this approach does justice neither to the level of sophistication of their interpretation and order-imposition (as noted above), nor to the force of Nietzsche's descriptions of the *ressentiment* subject's sickness- he seems to view them as essentially sick (GM III:14; GM III:15) – the use of the definite article in these passages 'the sick' [*die Kranken*] implies that this is their defining characteristic, and that there is no modicum of health present. This is further reinforced

[74] When Neuhouser surveys the *ressentiment* subject's 'illness' he points to four factors: 1) a measureless drive to make oneself suffer, 2) mendacity, 3) a lack of self-affirmation, and 4) a self-undermining dynamic present in the ascetic ideal. Without going into further discussion on these four characteristics, we may note that no reference is made here regarding the *ressentiment* subject's capacity to interpret and impose order upon encountered experiences.

by the fact that Nietzsche utilises a cautionary tone here, suggesting there is a danger of 'infection':

> The sick are the greatest danger for the healthy…. The sickly are the greatest danger to man: *not* the wicked, *not* the 'beasts of prey'. Those who, from the start, are the unfortunate, the downtrodden, the broken – these are the ones, the *weakest*, who most undermine life amongst men, who introduce the deadliest poison and scepticism into our trust in life, in man, in ourselves.
>
> (GM III:14)

Reducing health to interpretation therefore looks problematic when we consider this case in the *Genealogy.* It seems here that the *ressentiment* subjects meet the criterion of interpretation (even if they do not possess other marks of health). The reinterpretations they produce are sophisticated: they succeed in creating an entirely new valuation system which eventually supplants the previous moral system of the nobles. It is worth noting that the *ressentiment* subjects in fact display far more proficiency in interpretation than the nobles. As Nietzsche makes clear, the nobles are relatively unreflective agents and their health consists in their physicality:

> The chivalric-aristocratic value judgments are based on a powerful physicality, a blossoming, rich, even effervescent good health that includes the things needed to maintain it, war, adventure, hunting, dancing, jousting and everything else that contains strong, free, happy action.
>
> (GM I:7)

We get the sense that, as Ansell-Pearson suggests, the nobles merely live 'a life of immediate physical self-affirmation' (2007: xxi), unconstrained and uninhibited in their agency – there seems to be very little need for them to reinterpret and give meaning to any challenging experiences they encounter. For all intents and purposes, the nobles have free rein. The *ressentiment* subjects however are forced to utilise their capacity for interpretation. By Neuhouser's suggestion, this should mean that they are exhibiting an important mark of health yet Nietzsche is adamant that these individuals are pathological and sick. Although Nietzsche certainly praises the activity of (re)interpretation in his

works, (it is something which he himself professes to engage in), it seems that we cannot confine this activity to the healthy, given the cases cited in the *Genealogy.*

Let us take stock of what we have learnt so far about health with this approach of aligning health with broader Nietzschean concepts. It seems that this method of defining health has certain shortcomings: firstly, it can make it unclear exactly what work the concept health here is doing, and secondly, it often shows that the concept health and these other notions of overcoming and interpretation can come apart. These broader concepts receive attention on their own merit in Nietzsche's works, in varying guises, including contexts where references to health do not feature. This approach is perhaps a result of focusing on passages where Nietzsche discusses health in abstract terms (as in *The Gay Science* passages). We need only consult the distinct case-studies in which health is a clear concern, such as the slaves and the priests in the *Genealogy*, to see that we ought to be cautious about merely reducing health to other concepts, and that cases of health can be nuanced. This is not to say that the concept health cannot overlap with these other notions, but rather that we should note that it seems Nietzsche does not merely use the term 'health' as a substitute for these notions.

How else might we attempt to define Nietzschean health? Rather than reducing health to other concepts, which appears to be too broad an approach, we could refer directly to the mental apparatus in the self. Given that Nietzsche provides us with a detailed philosophical psychology, it seems that this would be a good starting point to think about health and to assess how this might be understood at the level of the activities of the drives and affects. One way of exploring this approach would be to adopt, as Huddleston (2017) calls it, a 'formal and dynamic approach' where the aim is to 'understand health in fundamentally descriptive (as opposed to normative) terms and to have the key descriptive properties be formal and dynamic ones concerning the strength and interrelation of the drives' (2017:151). Here the focus is on the arrangement of the drives within the self. For example, accounts might focus on the extent to which drives are unified or integrated, or whether there is a strong dominant master drive ruling over the hierarchy and so on.

However, Huddleston argues against defining health solely in formal and dynamic terms, suggesting that these accounts can at times be too simplistic and admit cases as 'healthy'

which Nietzsche himself emphasises are unhealthy, e.g. Socrates in *Twilight of the Idols.* In this example, (as Nietzsche characterises him) Socrates has a strong master drive (in this case reason) which strives to co-opts other drives towards its purpose: 'he is seemingly doing what the formal and dynamic account would prize as characteristic of health. Socrates has a strong master drive (viz., reason) which rules over his other drives' (2017:155). Given that in this work Nietzsche describes Socrates in very negative terms –presenting him as pathological and referring to reason as a 'tyrant' (TI 'Socrates' 11), the idea that according to the formal and dynamic account, Socrates is meeting a criterion of health, seems troublesome. So long as there is an adequately strong dominant master drive in the hierarchy, an individual can be termed healthy on these formal and dynamic accounts: 'one can live up to these standards, and still be pathological, in virtue of the quality of the drive that is dominant' (2017:156).

As a result of this concern, Huddleston suggests that we ought to supplement such accounts with normative considerations, i.e. to assess whether the *right* kind of drive is in charge: 'the core concept of Nietzschean health is not fully explicable except by reference to substantive normative terms. It matters not only that the drives are powerful and unified. It matters in addition what drives are regnant, and whether (by Nietzsche's lights) those drives are worthwhile and whether they ought to be regnant' (2017:136). We therefore ought to focus on 'the quality of particular drives in the self and the worth of their aims' (2017:136), and supplement any formal and dynamic accounts accordingly, otherwise we will fail to capture the central marks of Nietzschean health.

Huddleston emphasises the idea that it matters which drives exhibit strength and authority in the hierarchy– whether those drives are appropriate rulers within the self, and whether their aims are worthy. With regards to the example of Socrates, the problem here is that reason, according to Nietzsche, is not suited to be in authority of the hierarchy of drives:

> When Nietzsche claims that reason is made 'a tyrant' in the case of Socrates, the problem is not that reason is an *ineffective* ruler. It is that reason is an *inappropriate* ruler. The wrong drive is in a position of authority. Nefarious

> tyrants may get their way, but the important point is that they *shouldn't* get their way.
>
> (2017:156)

The fact that reason was strong and effective enough to overpower the 'anarchy of the instincts' in Socrates does not indicate that he was healthy: reason still remains an inappropriate ruler. Huddleston argues that this kind of normative judgement regarding whether the drive in charge *should* be in that position, is crucial if we wish to avoid miscategorising individuals as healthy or unhealthy. In maintaining that 'we do better to think of Nietzschean health in terms of a cluster of characteristic features' Huddleston seems to advocate acknowledging this normative dimension of the formal and dynamic accounts whilst also bearing in mind central marks of health such as 'resilience, discipline, vitality, a certain positive condition of the will to power, a certain tendency toward integration' (2017:136).

In this way, Huddleston argues that in order to begin to understand Nietzsche's concept of health, we need to move beyond a formal and dynamic account which is concerned with the arrangements of the drives and to instead consider the *normative* dimension of the formal and dynamic accounts, i.e. precisely which drives are in charge, and whether they are appropriate 'rulers'.

What does this account tell us about health? One consequence, is that there is an emphasis on the drive constitution of the individual rather than basic environmental factors,[75]and the implication seems to be that the health of an individual is relatively fixed. To go back to the example of Socrates: even if he places himself within a context conducive to his health (in *Ecce Homo*, Nietzsche discusses the importance of the appropriate nutrition and climate), and bearing in mind that what is healthy will vary according to different individuals, we get the sense that he cannot be classed as essentially healthy on this normative account in virtue of his drive constitution, i.e. an inappropriate ruler – reason-

[75] By basic environmental factors, I mean aspects like climate, food and so on, rather than social or cultural factors.

is in authority. Unless he changes his drive constitution in some significant way, Socrates will remain, in some sense, fundamentally sick: 'Socrates is among Nietzsche's paradigm cases of a sick, decadent person' (Huddleston: 2017:154).

If this sickness comes down to drive constitution, it is not obvious how one would go about changing or 'curing' themselves. Although Nietzsche gives some indication that we may try to deaden or eliminate drives (D 109; GS 290), it seems that a master drive will be especially strong and resistant to such procedures. Moreover, before attempting to deaden or eradicate a drive, there must be some recognition by the individual that this drive *ought* to be removed. Huddleston states that 'it matters whether the part (or drive) actually is *worth* retaining (in its present or transmuted form)' (2017:158). But this implies that the individual is able to assess the normative worth of their drives. We can envisage certain scenarios where this is feasible, for example an individual recognising that their drive to aggression is too prevalent and needs to be mitigated or removed (perhaps they come to realise this by observing external consequences, e.g. getting fired from their job due to being overly aggressive and confrontational in their interactions). In such a case, we can see how someone would be able to evaluate the worth of this drive and decide to work to remove it. In the case of Socrates, by contrast, it looks as though Socrates will deem a drive to reason as admirable and worth retaining. In this way, it seems that Socrates will remain 'sick' since there is an inability to self-correct, as it were. The possibility of amelioration or recovery from this sickness thus appears to be ruled out.

With regards to the slaves in the *Genealogy,* Huddleston's account is helpful in illuminating why they are classified as fundamentally sick. We get the sense that their sickness is due to their essential constitution, and that they are unable to do anything to alter this:

> Those who, from the start, are the unfortunate, the downtrodden, the broken – these are the ones, the weakest, who most undermine life amongst men, who introduce the deadliest poison and scepticism into our trust in life, in man, in ourselves. Where can we escape the surreptitious glance imparting a deep sadness, the backward glance of the born misfit revealing how such a man communes with himself, – that glance which is a sigh. 'If only I were some other

> person!' is what this glance sighs: 'but there's no hope of that. I am who I am: how could I get away from myself? And oh – I'm fed up with myself!' . . .
>
> (GM III:14)

So how exactly is the arrangement of their drives constituted? Given the palpability of emotions like vengefulness, envy, and hatred (see Chapter Four), it seems plausible to say that the drive to *ressentiment* (particularly in carrying out the so-called slave-revolt), dominates the drive hierarchy. *Ressentiment* is presented as characteristic of the slaves: Nietzsche repeatedly makes reference to 'the entrenched hatred and revenge of the powerless man' (GM I:10 cf. GM I:14; GM III:14), and ridicules the slaves for their inability to exercise this drive in physical form as the nobles would do: 'when *ressentiment* does occur in the noble man himself, it is consumed and exhausted in an immediate reaction, and therefore it does not poison'(GM I:10). The physical weakness of the slaves however, does not prevent their drive to revenge operating: they are able to exact an 'imaginary revenge' (GM I:10), which is achieved via self-deception. A drive to *ressentiment* is much maligned by Nietzsche- it is a mark of reactivity and weakness rather than activity and strength. The idea of this drive reigning in the drive hierarchy is therefore seen as a definitive indication of sickness.

The fact that this drive to *ressentiment* reigns in the hierarchy dictates that the slaves are sick no matter what other drive activity they display. Nietzsche presents the slaves' self-deception as the defining activity of these individuals: this phenomenon is central to the slave revolt in morality, and gives a good indication of what else is going on in their drive hierarchy. It is interesting to note that in generating their self-deceptions, the slaves in fact exhibit some admirable traits, specifically regarding their drive to artistic creation.

In Chapter Four, I argued that a key component of this 'sublime self-deception' is a pretense mechanism which is deployed in order to avoid cognitive dissonance (brought about by the conflict between the subjects' emotions of, for example, vengefulness, and their person-facing self-deceptive beliefs such as 'I am humble'). Such cognitive dissonance runs the risk of undermining the subject's self-deception. Intermediary states like pretense-states work to re-frame the subject's emotions via imagined narratives (cf. GM I:15), thus easing the cognitive dissonance that might be experienced. In this way, emotions like vengefulness are transferred to and expressed in an acceptable narrative

context – this emotional simulation which the narratives provide generate this 'imaginary revenge' (GM I:10), and allows their self-deceptive person-facing beliefs to remain intact. In other words, the subject is able to experience feelings of vengefulness whilst also still believing that they are meek and mild.

If, as I have argued these pretense-states occupy this role in the slave's mental economy, then this implies that the drive to creativity is particularly active and strong. From Nietzsche's descriptions of the slaves, he is careful to stress their creativity (GM I:10; GM III:15; GM III:20), and emphasises their imaginative capacity (see Chapter Four). The narratives they construct, as seen in the Tertullian passage (GM I: 15) are sophisticated, rich, and fantastical with vivid descriptions that are effective in provoking emotional responses. Such narratives require a high degree of artistry – these are not just simple stories which link up happenings in uninteresting ways. It therefore seems highly plausible that a drive to creativity is in operation here.

Huddleston has emphasised the idea the importance of the quality of drives and the worth of their aims. It seems like a drive to creativity fits this idea of an appropriate drive - Nietzsche is consistently committed to the notion of creativity across his works (e.g. BT, GS). Moreover, in GS 299 he is especially concerned with the question of 'how to make things beautiful, attractive, desirable for ourselves when they are not', i.e. what kind of aim this particular drive might have. He praises artists for teaching us the following lesson:

> Only they have taught us the art of viewing ourselves as heroes-from a distance and, as it were, simplified and transfigured-the art of staging and watching ourselves. Only in this way can we deal with some base details in ourselves.
>
> (GS 78)

What is interesting to note is that this seems to be precisely what the *ressentiment* subject is doing with his pretense-states. The narratives involve the subjects viewing themselves as the heroes of the story: 'he has conceived of the 'evil enemy', '*the evil one*' as a basic idea to which he now thinks up a copy and counterpart', the 'good one' – himself!' (GM I:10). Specifically, as I have discussed in Chapter Four, the *ressentiment* subject sees himself as the wronged hero, thus allowing these emotions of vengefulness – a 'base

detail' – to be validated. This is achieved through narratives, which stage these hero characters, simultaneously providing the *ressentiment* subjects with distance from their own realities and allowing them to engage with their emotions of vengefulness. It seems like this is a good candidate for what Nietzsche describes in GS 78. In this way, it looks as though the drive to creativity is playing a substantial role in the *ressentiment* subject's mental economy- it is called upon to generate these pretense-states, which work to safeguard their self-deceptions, and allow acceptable outlets for their 'base details.' Without the operations of this drive, the *ressentiment* subject might experience cognitive dissonance, and the self-deceptions which are required to enable them to endure their unfortunate conditions, are at risk of collapsing.

With this example of the pretense-states, at least, it seems that the slaves present a certain degree of commendable and perhaps even healthy drive activity: a valuable drive (the drive to creativity) with a worthy aim (reinterpreting or transfiguring certain aspects of the subject, i.e. the emotion of vengefulness) ranks highly in the hierarchy. Nevertheless, the fact that the drive to creativity has ultimately been co-opted by the reigning drive to *ressentiment* means that the slaves remain essentially sick.

This section has reviewed two different approaches to defining health. In doing so, I suggest that the following distinction emerges – there are ways in which an individual can persevere and continue to manage themselves/ function in light of difficult circumstances that do not always coincide with health. In other words, it seems that from the above examples, individuals can self-regulate even in sickness. This suggests that self-regulation may be a necessary condition for health but it is not a sufficient one. The next section will explore this distinction and attempt to outline some features of self-regulation.

5.2 Self-Regulation

This section aims to outline aspects of self-regulation, and to show how it can come apart from health. I will examine cases of self-regulation in both, according to Nietzsche, sick individuals, and in healthy individuals, identifying the key mechanisms involved.

An initial objection may be that the idea that someone self-regulates in sickness does not seem very plausible: initially it may seem difficult to put these things together – we may be accustomed to thinking that a healthy person is someone who is good at self-regulating, i.e. they are able to adapt quickly to their environment, and process experiences to maintain functional efficiency and overall physiological and/or psychological integrity.

However, a few examples may help show that it is not only healthy individuals who self-regulate, and that illness does not necessarily amount to an inability to engage in the process. Consider the case of a high-functioning alcoholic: the general consensus is that addiction is a form of illness, and that no matter how well-functioning the addict, this dependency is seen as a departure from health. Quite often, (high-functioning) alcoholics are able to maintain professional and personal lives, and do not use alcohol to achieve a state of inebriation, but rather to maintain a state of equilibrium. In other words, self-regulation manifests as a particular dependency, and the alcoholic drinks to maintain a psychological balance (and sometimes even a physiological balance – take, for example, delirium tremens which would mean that a surgeon would have to drink in order to avoid shaking hands). Other examples might be a patient suffering from motivation or defensive delusions, where these delusions arise as a result of the individual struggling to cope in some way, e.g. self-esteem issues, and cause the individual to view other people as responsible for their own difficulties. Or we might think of the confabulations of an amnesiac, who unwittingly falsifies memories in order to compensate for a lack of or a break in personal identity. With these examples, we can see that self-regulatory mechanisms, are evident in individuals who are in some way unwell or have impaired capacities.

I suggest that in the *Genealogy,* we find a clear case of self-regulation in sickness. As previously noted, Nietzsche is unambiguous about the state of the *ressentiment* subjects' health – they are *fundamentally* sick (GM I:6; GM III: 14, GM III:15). This illness is described in both physiological and 'spiritual' terms, and in passages like GM I:6 there is the suggestion that these two kinds of complaints can run alongside each other, or are perhaps interrelated:

> From the very beginning, there has been something unhealthy about these priestly aristocracies and in the customs dominant there, which are turned away from action and are partly brooding and partly emotionally explosive, resulting in the almost inevitable bowel complaints and neurasthenia which have plagued the clergy down the ages.
>
> (GM I:6)

This illness of the *ressentiment* subject is juxtaposed against the 'powerful physicality, blossoming, rich, even effervescent good health' of the nobles (GM I:7). Given Nietzsche's continued emphasis on the fact that the *ressentiment* subjects are essentially unhealthy, it seems vital that we incorporate this into any account of Nietzschean health, and accommodate the full force of these descriptions, i.e. moderating the category which the *ressentiment* subject falls under, from 'fundamentally sick' to 'poor levels of health' does not seem like an adequate move.

The category of *ressentiment* subject can be delineated further into two specific sub-types: the slaves and the priests. Both groups are, according to Nietzsche, essentially sick and both engage in self-regulatory processes. For the purposes of this chapter, however, I will focus on the self-regulation of the slaves. The reason for this is twofold. Firstly, Nietzsche's attitude towards the priest is somewhat ambivalent: at times, he even appears to admire the priest, referring to him as 'the direction-changer of *ressentiment*' (GM III:15), and 'this real artist in feelings of guilt' (GM III:20). Nietzsche acknowledges the ascetic priest's intellect: 'the greatest haters in world history, and the most intelligent [*die geistreichsten Hasser*], have always been priests: – nobody else's intelligence [*Geist*] stands a chance against the intelligence [*Geist*] of priestly revenge' (GM I: 7), and appears to be impressed by the priest's stalwart dedication to his ideals:

> The ascetic priest not only rests his faith in that ideal, but his will, his power, his interest as well. His right to exist stands and falls with that ideal: hardly surprising, then, that we encounter a formidable opponent in him, providing, of course, that we are opposed to that ideal?
>
> (GM III:11)

In this way, Nietzsche views the ascetic priest as a sophisticated paradigm of virile self-domination, since it is clear that the priest is seeking power and mastery over himself and others. The priest is evidently successful in self-regulating in that he persists and even 'thrives':

> Let us consider how regularly and universally the ascetic priest makes his appearance in almost any age; he does not belong to any race in particular; he thrives everywhere; he comes from every social class.
>
> (GM III:11)

Thus, despite being an 'apparent enemy of life, this negating one' the priest 'actually belongs to the really great *conserving* and *yes-creating* forces of life …' (GM III:13). Although the priest engages in self-regulatory processes, the fact that Nietzsche displays an ambiguous attitude towards this type of individual brings in certain considerations which may confuse the point.

Secondly, there is a tendency to view the slaves merely as tricked by the priest and thus as entirely passive individuals, blindly following the priest's teachings. However, as I hope to have shown in Chapter Four, the slaves in fact engage in sophisticated self-regulatory processes. Examining the slaves as a distinctive case of self-regulation will: (i) demonstrate that these individuals are more complicated than they may initially appear (in fact more complicated than the nobles), and (ii) bring out aspects of self-regulation we may not have noticed previously, showing crucial differences between this process and health.

As highlighted in the previous section, the reason the slaves are sick is due to the constitution of their drives and the fact that a drive to *ressentiment* is dominant. In this section, we will explore the consequences of this, and map out how this affects the mechanisms deployed to self-regulate.

The drive to *ressentiment* implies a certain dissatisfaction with oneself: as Huddleston notes 'the unhealthy person will be unable to bear certain parts of himself, in part because he has not (and cannot) sublimate them adequately into the whole that he potentially is. Because he is so distasteful to himself, he will need to resort to certain forms of self-deception' (2017:148). In GM III:14 Nietzsche makes it clear that the slaves display this

kind of attitude towards themselves: 'If only I were some other person!' is what this glance sighs: 'but there's no hope of that. I am who I am: how could I get away from myself? And oh – I'm fed up with myself!' …'

Here Huddleston implies that the activity of self-deception is not the preferred mechanism for processing content which the individual finds distasteful. If possible, this kind of material should either be incorporated or removed altogether. He refers to the following passage from *The Gay Science*:

> It is practiced by those who survey all the strengths and weaknesses of their nature and then fit them into an artistic plan until every one of them appears as art and reason and even weaknesses delight the eye. Here a large mass of second nature has been added; there a piece of original nature has been removed – both times through long practice and daily work at it. Here the ugly that could not be removed is concealed; there it has been reinterpreted and made sublime. Much that is vague and resisted shaping has been saved and exploited for distant views.
>
> (GS 290)

In this way, Nietzsche suggests that material ought to be either incorporated via reinterpretation such that it is made 'sublime' or eradicated from the self. The slaves' self-deception, as I have understood it in Chapter Four, allows for neither possibility.

Although, as I have suggested, this activity of self-deception can involve some admirable drive activity in that the drive to creativity is activated, this mechanism is not one which allows the subject to *incorporate* parts of himself into the 'whole that he potentially is'. Nietzsche suggests that this idea of incorporation or integration can be achieved by 'concealing' and 'reinterpreting'.

In Section One, when evaluating Neuhouser's suggestion that health involves interpretation, I argued that from the example of the slaves, it is clear that this activity does not only belong to the province of health, since the slaves are able to assign meaning to their suffering via the ascetic ideal with great success. These fundamentally sick individuals practise interpretation with great proficiency, and it seems that this is undertaken as a means of adapting to and managing their situation. Interpreting, and

assigning meaning and order to their experiences enables them to tolerate and endure their difficult circumstances.

When it comes to their own emotions of hatred, vengefulness, and envy, however, the matter is entirely different. Whilst the slaves are able to interpret and incorporate their suffering effectively, they are only able to temporarily mask their emotions – the distasteful parts of themselves - via pretense. Interpretation occurs in so far as these emotions are reframed via narratives and transferred to and expressed in this more acceptable context. But they are not reinterpreted or modified into emotions of a different kind, and it seems that this level of reinterpretation would have to occur for this material to be fully integrated. The idea that the slaves remain so unhappy with themselves, suggests that this incorporation has not taken place. Moreover, these pretense-states are only temporary: as I emphasised in Chapter Four, to say that the slaves continually engage in them would be tantamount to delusion, and there is no textual evidence in the *Genealogy* for this phenomenon. Any 'concealment' or 'reinterpretation' the pretense-states offer has a limited time-frame. I also argued that at no point is this negative affective content removed from the slaves' mental economy: if this were the case, they would not be as sick as Nietzsche suggests. Instead these emotions continue to lurk in the slaves' mental economy.

The mechanism of self-deception enables the slaves to endure and self-regulate, in order to manage their emotions and their situation, but it does not lead to incorporation of the parts of themselves they find distasteful.

Examining mechanisms of self-regulation in sick individuals brings the following features of self-regulation to light: (i) stability, endurance, or resilience, (ii) productivity and functional efficiency, and (iii) adaptivity. With the example of the slaves, their mechanism of pretense provides them not only with a means of tolerating their suffering, but also with a means of avoiding cognitive dissonance arising from tensions between their emotions and their self-deceptive beliefs. This ensures that the slaves are able to maintain a certain level of stability, in response to both their external circumstances, and the state of their inner world. The fact that the drive to creativity is implicated in this mechanism of self-deception means that the slaves, far from remaining inert, become very productive in generating these pretenses and their new values- these are created in an efficient

manner, and in this way at least, the slaves retain their functional efficiency. Finally, the slaves' self-deception emerges as a sign of adaptivity to unusual and difficult circumstances – their pretenses and the resulting new value system answer to clear aims, i.e. exacting revenge on the nobles. They are not created for random effect but are designed to be responsive to a particular situation.

In this way, self-regulation can occur very efficiently in cases where the individuals are sick. This is not a process which is restricted to the healthy as we may initially have thought. Moreover, it is evidence of the fact that individuals are to be conceived of as responsive and dynamic, and that even in sickness we possess 'the active capacity to create, to transform or give new forms, and meanings to things, including ourselves' (Janaway: 2017: 81). This also gives us some indication of why Nietzsche finds 'the sickly' in the *Genealogy* so alarming: 'the sick are the greatest danger for the healthy; harm comes to the strong not from the strongest but from the weakest. Do people realize this?' (GM III:15). We might normally conflate sickness with an inability to self-regulate and survive: the idea here is that if a particular group is sick, then they are weak, and if they are weak, then they will die out. With the example of the *ressentiment* subjects, however, Nietzsche is keen to emphasise that this is not necessarily the case. These individuals, according to Nietzsche, are certainly weak in terms of their norms. But they emerge as surprisingly resilient and able to survive/persist: through mechanisms like self-deception they are able to function efficiently. Moreover, the cautionary tone Nietzsche uses in places like GM III:15 implies that the sickly can infect and overthrow the healthy- we certainly see an instance of this in the idea that the morality of the nobles is displaced. The features of self-regulation which the slaves exhibit – stability, endurance or resilience, productivity and functional efficiency, and adaptivity are enough to ensure that these individuals survive.

However, as noted in the example of the slaves, one further aspect of self-regulation is missing when we consider sick individuals - the idea of incorporation and psychological integrity. Nietzsche alludes to at least three methods by which incorporation can take place: via forgetting, through the use of (autobiographical) memory, and through the process of spiritualisation.

In Chapters Two and Three, I discussed the importance of forgetting and autobiographical memory as mechanisms of self-regulation. Forgetting allows the individual to 'digest' experiences and to dial down the negative affective valence associated with difficult experiences. In this way, space is cleared in the mental economy, and the health of the self is not only preserved but bolstered. Experiences are incorporated in so far as the useful components are taken in and made use of (as in the process of digestion and uptake of nutrients/expulsion of toxins). Autobiographical memory allows for events to be included in the narrative self, and in Chapter Three, I suggested that we can view Nietzsche constituting such a self through autobiographical memory in *Ecce Homo*. The interactions between the drives and autobiographical memory allow material to be incorporated into an emerging narrative which bolsters the well-being of the individual. In both cases, I have argued that neither forgetting nor the construction of a narrative self through autobiographical memory is to be understood as a form of whitewashing or censorship, but rather as a means of incorporation and integration of experiences into the self. Huddleston (2017) and Gemes (2009b) refer to a further method of integration or incorporation - Nietzsche's notion of spiritualisation (TI 'Morality' 3): here drives take on more refined, acceptable forms under the influence of other, stronger drives. The notion of incorporation and the ensuing psychological integrity seems pertinent for what Nietzsche describes in GS 290 and is one aspect of self-regulation which does appear to coincide with health.

Commentators such as Reginster (1997) have drawn attention to the importance of integrity within the self. Reginster argues that Nietzsche aligns this notion of integrity qua psychic harmony with nobility:

> The early noble's ''predominance did not lie mainly in physical strength but in strength of soul-they were more whole human beings" (BGE, 257). Strength of soul" and "wholeness" evoke a certain notion of integrity, which therefore stands out as a crucial feature of nobility of character.
>
> (1997: 298)

Gemes (2009b) and Katsafanas (2011) also highlight the importance of integrity within the self (though they characterise this in different ways: Gemes understands integrity to consist in one master drive uniting and ruling over others; Katsafanas takes integrity to

consist in a unity between reflective and unreflective components in the mental economy). It seems like we can understand psychological integrity as a mark of health in that the individual does not experience any cognitive dissonance – this can be detrimental to a sense of self and agency, and well-being in general.

The below table illustrates what we have learnt about self-regulation from the examples in this chapter.

Self-Regulation		
General Methods	Key Characteristics	Psychological Mechanisms involved
Interpretation	Stability/Endurance/Resilience	Forgetting
Incorporation	Productivity and Functional Efficiency	Autobiographical Memory
Extirpation	Adaptivity	Self-Deception (Pretense Mechanisms)
	Physiological/Psychological Integrity	Spiritualisation

Self-regulation is a phenomenon which individuals can engage in to varying degrees. At a very basic level, it can involve our immediate contact with the world: we might think of Nietzsche's example of forgetting as perceptual filtering with the example of the leaf in 'On Truth and Lies'. Here, too much perceptual data would be overwhelming for an individual and would not be processed in their mental economy. At a more sophisticated level, self-regulation can involve the processing of affective experiences (as we have seen in Chapters Two, Three, and Four), and this kind of self-regulation can occur in radically different ways as the above table illustrates. Individuals can self-regulate in this manner

in more or less successful ways – success here can be understood as tantamount to survival.

With regards to the relationship between health and self-regulation, the picture is complicated. The use of forgetting, autobiographical memory, and the process of spiritualisation are all generally aligned with health, and draw on each of the methods above: interpretation (e.g. autobiographical memory), extirpation (e.g. forgetting), and incorporation (e.g. spiritualisation). I have argued that self-deception is presented as a characteristic of sickness, but still very much belongs to the phenomenon of self-regulation. The fact that the slaves survive through this successful mode of self-regulation has nothing to do with health.

We have seen how self-regulation and health can come apart. Just because one is sick, does not mean that one does not self-regulate efficiently. We have seen this with the slaves in the *Genealogy*. On the flipside, just because one is healthy does not mean that one is self-regulating efficiently. An interesting example for the latter case is the nobles in the *Genealogy*. Nietzsche emphasises that these individuals are healthy (GM I:7; GM II:1) and through their capacity to forget they self-regulate. Nietzsche praises their use of this capacity to forget, and the implication is that this affords them with a sense of well-being, self-esteem, and psychological integrity in general as they are not held back by negative affective material: 'a man like this shakes from him, with one shrug, many worms which would have burrowed into another man' (GM I:10). However, the fact that the nobles seem to rely on this particular method of self-regulation leaves them vulnerable: when confronted with the slave revolt, this group's resilience and ability to adapt and survive is tested and found wanting. Forgetting does not appear to be particularly useful in dealing with the newly minted values of slave-morality and the intellectual agon of the priests, and the fact that the slave revolt exploits aspects of the nobles such as their drives to competition, or domination. There does not seem to be anything to 'forget' in particular.

We might speculate that a potential means of combatting the slave revolt would be via further reinterpretation but it does not seem that the nobles opt for this: they do not create new values, reinterpret the slaves' values, or engage in any kind of overhaul of their own noble values. As a result, they eventually succumb to the slave revolt in that

their morality is displaced by the slave-morality. In this way, the nobles' self-regulation is somewhat fragile: they are unable to adapt even though they are 'healthy', since interpretation does not appear to be in their arsenal (Nietzsche seems to assign this capacity of interpretation to artistic or intellectual types- they nobles of the *Genealogy* are not presented as falling into either category). With this example of the nobles, we can see that interpretation is not always necessary for health, and that health does not always amount to resilience.

I will conclude this section by briefly reflecting on the relationship between the concepts of self-regulation and health, and Nietzsche's notion of life. Much like health, life is a concept to which Nietzsche frequently refers but never really clarifies. It is often presented as a normative concept like flourishing, and commentators (e.g. Neuhouser) have tended to align it with power. It may seem like a truism that health and life go hand but as we will see in this final example, this is not necessarily the case. In the third essay of the *Genealogy*, Nietzsche outlines the nature and function of the ascetic ideal of the priest and the slaves. Initially, as an ideal of self-abnegation and self-denial, the ascetic ideal may seem to operate against life. However, as Nietzsche makes clear, this particular ideal is 'a trick for the preservation of life' (GM III:13) and the priest is in fact, in the service of life:

> You take my meaning already: this ascetic priest, this apparent enemy of life, this *negating one*, – he actually belongs to the really great *conserving* and *yes-creating* forces of life . . . His 'no' that he says to life brings a wealth of more tender 'yeses' [*eine Fülle zarterer Ja's* (sic)] to light as though by magic; and even when he *wounds* himself, this master of destruction, self-destruction, – afterwards it is the wound itself that forces him to *live* ...
>
> (GM III:13)

By utilising the ascetic ideal to propagate ideas of sin and guilt and thus provide meaning to the slaves' suffering, the priest enables the slaves to continue living and directs them to 'self-discipline, self-surveillance, and self-overcoming' (GM III:16). Chapter Four discussed how the priest's teachings inspire the pretense mechanisms used by the slaves in their self-deception, and that this emerges as a successful means of self-regulation. In this way, Nietzsche suggests that the activity of the priest arises from 'the protective and

healing instincts of a degenerating life, which uses every means to maintain itself and struggles for its existence' (GM III:13). Nietzsche presents the priest as a resilient type: as previously noted, the priest self-regulates very efficiently or 'thrives' cf. GM III:11. Given that the priest is fundamentally sick however, we now have an example in which both life and self-regulation are present, but health is distinctly lacking. By setting life and health against the background of self-regulation, we can see how complicated the relationship between these concepts can be and that they do not necessarily map onto each other as we may initially assume.

Conclusion

In summary, by triangulating the notion of self-regulation with Nietzsche's remarks on health, and the case-studies of health and sickness he provides in works like the *Genealogy*, we have been able to uncover key methods and characteristics of the self-regulation process. Self-regulation emerges as a distinctive phenomenon in its own right which does not always coincide with instances of health. In this way, it appears as a necessary but not sufficient condition for health. We have also been able to establish that, when viewing individuals through this lens of self-regulation, Nietzsche's concepts of life and health can also be independent of each other in cases like that of the ascetic priest and the slaves. The 'instinct for self-restoration' to which Nietzsche refers in *Ecce Homo* is therefore not restricted to the 'higher' types but appears to be more widespread and persists even in 'sick' individuals.

Conclusion

With this thesis, I hope to have proposed a new framework in which to explore some of Nietzsche's remarks on the self. I have characterised self-regulation as a phenomenon exhibited by all individuals, in various ways, and to differing degrees of efficiency and success, where success is understood as allowing the individual to survive, (and, in some cases, to flourish or prosper).

Self-regulation has been understood as the maintenance of the individual's psychological integrity in both a basic sense, i.e. everyday processing of experiences, and in more extreme cases, when the subject finds themselves in challenging or adverse circumstances. It is interesting to note that this process of self-regulation does not always coincide with health, as we may have originally assumed.

I have also investigated how the process of self-regulation may provide the necessary conditions for individuals who (in Nietzsche's eyes), are capable of the 'higher' activity of self-formation. We have seen that incorporation, in particular, emerges as a particular mode of self-regulation which would play a part in self-formation.

I have examined components of the Nietzschean self beyond the drives and affects, how their functional profiles differ from the drives and affects, and the way in which these components interact with them. In this way, we have seen some of the richness and nuance of the philosophical psychology envisaged by Nietzsche.

To close, I will offer some further lines of enquiry which this project has raised.

So far, I have focused on examining processes which are not restricted to those whom Nietzsche terms 'higher types'. I have looked at processes which can occur in ordinary individuals, and in those whom Nietzsche views as 'weak' or even 'sick'. But what would self-regulation look like in Nietzsche's higher, artistic individuals, for example Goethe? How might this differ from the ordinary case? The particular drive constitution in such types might dictate specific modes of self-regulation, beyond those explored in the thesis.

A second, and related enquiry would be to ascertain what a sufficient condition(s) for self-formation would be in individuals who successfully self-regulate, and who possess

the drive constitution required for self-formation. I have suggested that self-regulation is a necessary but not sufficient condition for the activity of self-formation. Further examination between higher level processes like artistic creation, and more basic regulatory procedures, may shed some light on what these sufficient conditions for self-formation might be.

Lastly, throughout the chapters, the term 'capacity' has been used. Forgetting, memory, and imagination (used in generating pretenses) have all been characterised as capacities. A further step would be to explore whether we can arrive at a unified notion of a capacity (as has been attempted with the notion of a 'drive' or 'affect'), and to examine how such a structure would interact with other components of the self such as conscious willing, or the intellect.

I hope to have shown ways in which we may supplement drive-theory accounts, and to suggest that the components of a Nietzschean self examined in this thesis illuminate how such a self behaves and processes experiences.

Bibliography

Writings by Nietzsche

Reference edition of Nietzsche's works:

Werke: Kritische Gesamtausgabe, Colli, G., and Montinari, M., (eds.), Berlin: de Gruyter, 1967 ff.

Frühe Schriften, Mette, H. J., Schlechta, K., and Koch, C., (eds.), Munich: Beck, 1994.

Sämtliche Werke: Kritische Studienausgabe, Colli, G., and Montinari, M., (eds.), Berlin: de Gruyter, 1980 ff.

Translated Editions of Nietzsche's works:

The Birth of Tragedy and Other Writings, Speirs, R., (trans.), Geuss. R., and Speirs, R., (eds.), Cambridge: Cambridge University Press, 1999 (1872).

Untimely Meditations, Hollingdale, R.J., (trans.), Cambridge: Cambridge University Press, 1983 (1873–6).

Human, all too Human: a Book for Free-Spirits, Hollingdale, R.J., (trans.) Cambridge: Cambridge University Press, 1986 (Vol. I, 1878; Vol. II, 1879–80).

Daybreak, Hollingdale, R.J., (trans.), Cambridge University Press, 1982 (1881).

The Gay Science, Kaufmann, W., (trans.) New York: Vintage, 1974 (1887).

Thus Spoke Zarathustra, Del Caro, A., (trans.), Del Caro, A., and Pippin, R., (eds.) Cambridge: Cambridge University Press 2006 (1883–5).

Beyond Good and Evil: Prelude to a Philosophy of the Future, Hollingdale, R.J., (trans.), Penguin: 2003 (1886).

On the Genealogy of Morality, Diethe, C., (trans.) Cambridge University Press, 2010 (1967).

The Anti-Christ, Ecce Homo, Twilight of the Idols and Other Writings, Norman. J., (trans.), Ridley, A., and Norman, J., (eds.) Cambridge: Cambridge University Press, 2005 (1895, 1908, 1888).

The Wagner Case, Kaufmann, W., (trans.) New York: Vintage, 1966 (1888).

The Will to Power, Kaufmann, W., and Hollingdale, R.J., (trans.), Kaufmann, W., (ed.), New York: Vintage, 1967 (1901, 1906).

Secondary Literature

Acampora, C, D., (2006), *Nietzsche's On the Genealogy of Morals: Critical Essays,* USA: Rowman & Littlefield Publishers inc.

Alfano, M., (2013), 'The Most Agreeable of all Vices: Nietzsche as a Virtue Epistemologist' in the *British Journal for the History of Philosophy,* 21 (4):767-790.

——(forthcoming), 'A Schooling in Contempt: Emotions and the Pathos of Distance' in Katsafanas, P., (ed.), *The Nietzschean Mind (Routledge Philosophical Minds),* Routledge.

Anderson, M. C., and Hanslmayr, S., (2014) 'Neural Mechanisms of Motivated Forgetting' in *Trends in Cognitive Sciences,* 18(6):279-292.

Anderson, M. C., and Huddleston, E., (2012), 'Towards a Cognitive and Neurobiological Model of Motivated Forgetting' in Belli, R.F., (ed.), *True and False Recovered Memories: Toward a Reconciliation of the Debate,* Springer Science+Business Media.

Anderson, R. L., (1998), 'Truth and Objectivity in Perspectivism' in *Synthese*, 115 (1):1-32.

Anderson, R. L., (2005), 'Nietzsche on Truth, Illusion, and Redemption' in *European Journal of Philosophy*, 13 (2):185–225.

Anderson, R. L., (2012), 'What is a Nietzschean Self?' in Janaway, C., and Robertson, S., *Nietzsche, Naturalism, and Normativity,* Oxford: Oxford University Press.

Babel, P., (2015), 'Memory of Pain induced by Physical Exercise' in *Memory* 24(4):548-59.

Bayne, T., and Fernández, J., (2009), *Delusion and Self-Deception: Affective and Motivational Influences on Belief Formation*, New York: Psychology Press.

Bermudez, J, L., (2017), 'Memory and Self-Consciousness' in Bernecker, S., and Michaelian, K., *The Routledge Handbook of Philosophy of Memory*, Oxford: Routledge.

Blustein, J., (2008), *The Moral Demands of Memory*, Cambridge: Cambridge University Press.

—— (2014), *Forgiveness and Remembrance: Remembering Wrongdoing in Personal and Public Life*, Oxford: Oxford University Press.

Borges, L, J., (1941), 'Funes el memorioso' in *Ficciones,* (Madrid: Alianza Editorial).

Conway, M, A., (2005), 'Memory and the Self' in *Journal of Memory and Language*, 53(4): 594–628.

Conway, M. A., Singer, J. A. and Tagini, A. (2004), 'The Self and Autobiographical Memory: Correspondence and Coherence' in *Social Cognition*, 22(5), 491-529.

Conway, D, W., (1998), *Nietzsche: On Morality,* London: Routledge.

Davidson, D., (1982), 'Paradoxes of Irrationality' in *Philosophical Essays on Freud*, Wollheim, R., and Hopkins, J., (eds.), Cambridge: Cambridge University Press.

—— (1985), 'Deception and Division' in *Actions and Events*, LePore, E., and McLaughlin, B., New York: Basil Blackwell.

Deleuze, G., *Nietzsche & Philosophy*, Tomlinson, H., (trans.) New York: Columbia University Press, 1983 (1962).

Dennett, D., (1992), 'The Self as a Center of Narrative Gravity' in *Consciousness and Self: Multiple Perspectives*, Kessel, F., Cole, P., and Johnson, D., (eds.), Hillsdale, NJ: L. Erlbaum.

Elgat, G., (2015), 'Slave Revolt, Deflated Self-Deception' in *British Journal for the History of Philosophy,* 23 (3):524-544.

Gardiner, J, M., (2001), 'Episodic Memory and Autonoetic Consciousness: A First-Person Approach' in *Philosophical Transactions of the Royal Society B: Biological Sciences,* 356(1413): 1351–1361.

Gardner, S., (2009), 'Nietzsche, The Self, and The Disunity of Philosophical Reason' in Gemes, K., and May, S., *Nietzsche on Freedom and Autonomy,* Oxford: Oxford University Press.

——(2015), 'Nietzsche and Freud: The 'I' and its drives' in Ryan, B., Mayer Branco, M, J., and Constancio, J., (eds.), *Nietzsche and the Problem of Subjectivity*, Berlin: De Gruyter.

Gemes, K., (2009a), 'Nietzsche on Free Will, Autonomy, and the Sovereign Individual' in Janaway, C., and Robertson, S., *Nietzsche, Naturalism, and Normativity,* Oxford: Oxford University Press.

—— (2009b) "Freud and Nietzsche on Sublimation." *Journal of Nietzsche Studies* 38: 38–59.

Gendler, T, S., (2007), 'Self-Deception as Pretense' in *Philosophical Perspectives,* 21 (1):231 – 258.

——(2008), 'Alief and Belief' in *Journal of Philosophy,* 105 (10):634-663.

——(2010), *Intuition, Imagination, and Philosophical Methodology*, Oxford: Oxford University Press.

Goldie, P., (2012a), *The Mess Inside: Narrative, Emotion, and the Mind*, Oxford: Oxford University Press.

Goldie, P., (2012b), 'The narrative sense of self' in *Journal of Evaluation in Clinical Practice,* 18: 1064–1069.

Guay, R., (2006), 'The 'I's Have it: Nietzsche on Subjectivity' in *Inquiry* 49 (3): 218-241.

Hales, S., and Welshon, R., (2000), *Nietzsche's Perspectivism,* University of Illinois Press.

Huddleston, A., (2017), 'Nietzsche on the Health of the Soul' in *Inquiry: An Interdisciplinary Journal of Philosophy* 60 (1-2):135-164.

Hutto, D, D., (2017), 'Memory and Narrativity' in Bernecker, S., and Michaelian, K., *The Routledge Handbook of Philosophy of Memory*, Oxford: Routledge.

James, W., (1890), *The Principles of Psychology,* Vol.1, New York: Dover Publications, Inc.

Janaway, C., (2007), *Beyond Selflessness: Reading Nietzsche's Genealogy,* Oxford: Oxford University Press.

—— (2009), 'Autonomy, Affect, and the Self in Nietzsche's Project of Genealogy' in Gemes, K., and May, S., *Nietzsche on Freedom and Autonomy,* Oxford: Oxford University Press.

—— (2012), 'Nietzsche on Morality, Drives, and Human Greatness' in Janaway, C., and Robertson, S., *Nietzsche, Naturalism, and Normativity,* Oxford: Oxford University Press.

——(2017), 'Attitudes to Suffering: Parfit and Nietzsche' in *Inquiry: An Interdisciplinary Journal of Philosophy,* 60 (1-2):66-95.

Jensen, A, K., (2013), *Nietzsche's Philosophy of History,* Cambridge: Cambridge University Press.

Katsafanas, P., (2005) 'Nietzsche's Theory of Mind: Consciousness and Conceptualisation' in *European Journal of Philosophy,* 13 (1):1-31.

——(2011), 'The Concept of Unified Agency in Nietzsche, Plato, and Schiller' in *Journal of the History of Philosophy,* 49 (1): 87-113.

—— (2013), 'Nietzsche's Philosophical Psychology' in Richardson, J., and Gemes, K., (eds.) *The Oxford Handbook of Nietzsche,* Oxford: Oxford University Press.

—— (2014) 'Nietzsche and Kant on the Will: Two Models of Reflective Agency' in *Philosophy and Phenomenological Research,* 89 (1):185-216.

——(2015), 'Kant and Nietzsche on Self-Knowledge' in Ryan, B., Mayer Branco, M, J., and Constancio, J., (eds.), *Nietzsche and the Problem of Subjectivity*, Berlin: De Gruyter.

——(2016), *The Nietzschean Self: Moral Psychology, Agency, and the Unconscious*, Oxford: Oxford University Press.

Klein, S. B., Loftus, J., and Kihlstrom, J. F., (2002, 'Memory and Temporal Experience: The Effects of Episodic Memory Loss on an Amnesiac Patient's Ability to Remember the Past and Imagine the Future' in *Social Cognition,* 20 (5): 353-79.

Klein, S, B., and Nichols, S., (2012), 'Memory and the Sense of Personal Identity' in *Mind*, 121(483): 677–702.

Klein, S. B, (2013), 'Making the case that episodic recollection is attributable to operations occurring retrieval rather than to content stored in a dedicated subsystem of long-term memory' in *Frontiers in Behavioral Neuroscience* 7 (3):1-14.

Korsgaard, C., (2009), *Self-Constitution: Agency, Identity, and Integrity,* Oxford: Oxford University Press.

Leiter, B., (1998), 'The Paradox of Fatalism and Self-Creation in Nietzsche' in Christopher Janaway (ed.), *Willing and Nothingness: Schopenhauer as Nietzsche's Educator* Oxford: Clarendon Press.

——(2002) *Nietzsche on Morality* (London: Routledge).

—— (2007) 'Nietzsche's Theory of the Will' in *Philosophers' Imprint*, vol.7: no.7.

—— (2010) 'Gardner on "Nietzsche, the Self, and the Disunity of Philosophical Reason"', Brian Leiter's Nietzsche Blog, URL=

http://brianleiternietzsche.blogspot.co.uk/2010_08_01_archive.html.

—— (2013) 'Nietzsche's Moral and Political Philosophy' *The Stanford Encyclopaedia of Philosophy,* (Spring 2013 Edition), Edward N. Zalta (ed.), URL=

http://plato.stanford.edu/archives/spr2013/entries/nietzsche-moral-political/.

Letteri, M, R., (1990), 'The Theme of Health in Nietzsche's Thought' in *Man and World,* 23 (4): 405-417.

Marshall, J., (2011) 'Forgetfulness is Key to a Healthy Mind' in *New Scientist*.

Martin, C, W., (2006), 'Borges forgets Nietzsche' in *Philosophy and Literature* 30.1 (2006) 265-276.

——(2009), *The Philosophy of Deception*, New York: Oxford University Press.

May, S., (2009), 'Nietzsche and the Free Self' in Gemes, K., and May, S., *Nietzsche on Freedom and Autonomy,* Oxford: Oxford University Press.

McLaughlin, B. P., and Rorty, A. O., (eds.), (1988), *Perspectives on Self-Deception,* Berkeley: University of California Press.

Mele, A.R., (2000), *Self-Deception Unmasked*, Princeton: Princeton University Press.

—— (2003), 'Emotion and Desire in Self-Deception' in Hatzimoysis, A., (ed.), *Philosophy and the Emotions,* Cambridge: Cambridge University Press.

Moran, R., (1994), 'The Expression of Feeling in Imagination' in *Philosophical Review,* 103 (1):75-106.

Nehamas, A., (1985), *Nietzsche: Life as Literature*, Cambridge, MA: Harvard University Press.

Neuhouser, F., (2014), 'Nietzsche on Spiritual Health and its Promise' in *Journal of Nietzsche Studies*, 45 (3):293-314.

—— (forthcoming), 'Nietzsche on Spiritual Health and Cultural Pathology' in Katsafanas, P., (ed.), *The Nietzschean Mind (Routledge Philosophical Minds),* Routledge.

Noordhof, P., (2009), 'The Essential Instability of Self-Deception' in *Social Theory and Practice*, 35(1): 45–71.

Parker, E. S., Cahill, L., and McGaugh, J.L., (2006), 'A Case of Unusual Autobiographical Remembering' in *Neurocase* 12 (1): 35-49.

Pears, D., (1984), *Motivated Irrationality*, New York: Oxford University Press.

—— (1991), 'Self-Deceptive Belief Formation,' in *Synthese*, 89: 393–405.

Perry, J., (1979), 'The Problem of the Essential Indexical' in *Noûs*, 13 (1): 3-21.

—— (2007),'"Borges and I" and "I"', The Amherst Lecture in Philosophy 2:1-16.

Poellner, P., (1995), *Nietzsche and Metaphysics,* Oxford: Clarendon Press.

——(2004) ‘‘Self-Deception, Consciousness and Value’ in *Journal of Consciousness Studies,* 11 (10-11): 44-65.

——(2015) ‘*Ressentiment* and the Possibility of Intentional Self-Deception’ in Dries, M., and Kail, P., *Nietzsche on Mind and Nature,* Oxford: Oxford University Press.

Porcher, J. E., (2012), ‘Against the Deflationary Account of Self-Deception’ in *Humana Mente,* (20):67-84.

——(2014), ‘Is Self-Deception Pretense?’ in *Manuscrito*, 37 (2):291-332.

Reginster, B., ‘Nietzsche on *Ressentiment* and Valuation’ in *Philosophy and Phenomenological Research,* 57 (2): 281-305.

Riccardi, M., (2011), ‘Nietzsche’s Sensualism’ in *European Journal of Philosophy,* 21 (2):219-257.

—— (2015), ‘Nietzsche on the Embodiment of the Mind and Self’ in Ryan, B., Mayer Branco, M, J., and Constancio, J., (eds.), *Nietzsche and the Problem of Subjectivity*, Berlin: De Gruyter.

——(forthcoming), ‘A Tale of Two Selves: Nietzsche and the Contemporary Debates on the Self’ in Katsafanas, P., (ed.), *The Nietzschean Mind (Routledge Philosophical Minds),* Routledge.

Richardson, J and Leiter, B., (eds.), (2001) *Nietzsche*, New York: Oxford University Press.

Richardson, J., (2008), ‘Nietzsche’s Problem of the Past’ in Dries, M., (ed.), *Nietzsche on Time and History,* Berlin: De Gruyter.

Roache, R., (2016), ‘Memory and Mineness in Personal Identity’ in *Philosophical Psychology* 29 (4): 479-489.

Rutherford, D., (forthcoming), ‘Nietzsche and the Self’ in Katsafanas, P., (ed.), *The Nietzschean Mind (Routledge Philosophical Minds),* Routledge.

Schechtman, M., (1994), ‘The Truth About Memory’ in *Philosophical Psychology*, 7(1): 3–18.

——(1996), *The Constitution of Selves,* Ithaca: Cornell University Press.

——(2005), 'Personal Identity and the Past' in *Philosophy, Psychiatry, and Psychology*, 12 (1):9-22.

——(2011), 'Memory and Identity' in *Philosophical Studies*, 153(1): 65–79.

——(2011), 'The Narrative Self' in Gallagher, S., (ed.), *The Oxford Handbook of the Self*, Oxford: Oxford University Press.

Shoemaker, S., (1996), *The First-Person Perspective and Other Essays*, Cambridge: Cambridge University Press.

Stern, T., (2015) 'Against Nietzsche's "Theory of Drives"' in *Journal of the American Philosophical Association*, 1 (1):121—140.

Sutin, A, R., and Robins, R, W., (2008), 'When the "I" looks at the "Me": Autobiographical Memory, Visual Perspective, and the Self' in *Consciousness and Cognition*, 17(4): 1386–1397.

Swanton, C., (2015), *The Virtue Ethics of Hume and Nietzsche*, Wiley Blackwell.

Tubert, A., (forthcoming), 'Nietzsche and Self-Constitution' in Katsafanas, P., (ed.), *The Nietzschean Mind (Routledge Philosophical Minds),* Routledge.

Teroni, F., (2017), 'The Phenomenology of Memory' in Bernecker, S., and Michaelian, K., *The Routledge Handbook of Philosophy of Memory*, Oxford: Routledge.

Watson, L, A., and Berntsen, D., (eds.) (2015), *Clinical Perspectives on Autobiographical Memory,* Cambridge: Cambridge University Press.

Welshon, R., (1992), 'Nietzsche's Peculiar Virtues and the Health of the Soul' in *International Studies in Philosophy* 24 (2): 77–89.

Williams, B. (1994), 'Nietzsche's Minimalist Moral Psychology' in R. Schacht (ed.), *Nietzsche*, *Genealogy, Morality,* Berkeley: University of California Press.

Wilson, A, E., and Ross, M., (2003), 'The Identity Function of Autobiographical Memory: Time is on our Side' in *Memory* 11 (2): 137-49.

Zahavi, D., (2007), 'Subjectivity and the First-Person Perspective' in *Southern Journal of Philosophy* 45 (S1):66-84.